②

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

DECISION MAKING HEURISTICS AND BIASES IN SOFTWARE PROJECT MANAGEMENT: AN EXPERIMENTAL INVESTIGATION

by

Daniel Ronan

March 1990

Thesis Advisor Tarek K. Abdel-Hamid

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Postgraduate School; 6b. OFFICE SYMBOL (If applicable) Code 37	7a. NAME OF MONITORING ORGANIZATION Naval Postgraduate School
6c. ADDRESS (City, State, and ZIP Code) Monterey, California 93943-5000	7b. ADDRESS (City, State, and ZIP Code) Monterey, California 93943-5000
8a. NAME OF FUNDING/SPONSORING ORGANIZATION; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
DECISION MAKING HEURISTICS AND BIASES IN SOFTWARE PROJECT MANAGEMENT: AN EXPERIMENTAL INVESTIGATION

12. PERSONAL AUTHOR(S)
Ronan, Daniel

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Master's Thesis	FROM ____ TO ____	1990, March	112

16. SUPPLEMENTARY NOTATION
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Heuristics; Biases; Software Project Management; Anchoring; Social Loafing

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Software project development has been plagued with an infamous reputation for cost overruns, late deliveries, poor reliability and users' dissatisfaction. Much of this blame has been placed on the managerial side of software development. The Systems Dynamic Model of Software Project Management is a quantitative model of software project dynamics that is attempting to gain some valuable insight into the managerial side of developing software systems.

The objective of this thesis is to use the Systems Dynamic Model's gaming interface to investigate managerial heuristics and biases in software project management. Specifically, three experiments were executed to determine the effect of "anchoring" on productivity estimation, the effect of poor cost estimation on staffing decisions and the effect of "social loafing" on a software project's staffing decisions, final cost and final duration.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Prof. Tarek K. Abdel-Hamid	22b. TELEPHONE (Include Area Code) (408) 646-2686; 22c. OFFICE SYMBOL Code AS/Ah

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete

Approved for public release; distribution is unlimited

Decision Making Heuristics and Biases in Software Project Management: An Experimental Investigation

by

Daniel Ronan
Lieutenant, United States Coast Guard
B.S., United States Coast Guard Academy, 1984

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN INFORMATION SYSTEMS

from the

NAVAL POSTGRADUATE SCHOOL
March 1990

Author: Daniel Ronan

Approved by: Tarek K. Abdel-Hamid, Thesis Advisor

Nancy C. Roberts, Second Reader

David R. Whipple, Chairman
Department of Administrative Sciences

ABSTRACT

Software project development has been plagued with an infamous reputation for cost overruns, late deliveries, poor reliability and users' dissatisfaction. Much of this blame has been placed on the managerial side of software development. The Systems Dynamic Model of Software project Management is a quantitative model of software project dynamics that is attempting to gain some valuable insight into the managerial side of developing software systems.

The objective of this thesis is to use the Systems Dynamic Model's gaming interface to investigate managerial heuristics and biases in software project management. Specifically, three experiments were executed to determine the effect of "anchoring" on productivity estimation, the effect of poor cost estimation on staffing decisions and the effect of "social loafing" on a software project's staffing decisions, final cost and final duration.

TABLE OF CONTENTS

I. INTRODUCTION

A. BACKGROUND

The rapidly changing technology of the past few years has driven the cost of computer hardware lower and lower. With every drop in hardware price and increase in power, a rising number of users are demanding additional and more complex software. Although the improvements in hardware performance have been dramatic, the improvements in software productivity have only increased at a sluggish four percent annual rate. This slow rate coupled with the fact that computer programmers and project managers are in short supply has created a logjam of software applications waiting to be designed and coded. Those that finally make it through the bottleneck and get developed are, with all too much frequency, unreliable and/or mired in cost and schedule overruns. [Ref. 1:pp. 100-101]

In a recent article, Brenton R. Schlender places the majority of the blame for the software crisis on the management side of the software industry. He states,

> ...the biggest obstacles to effective, economical software development are managerial. In case after case, the cause of delayed or botched software invariably boils down to bad planning, organizational rivalries, unrealistic scheduling, or the inability of techies to grasp the business problems they are trying to solve. [Ref. 1:p. 107]

The solution for the software crisis, or at least a remedy to release some of its stranglehold on the software industry, appears to lie on the managerial side of software development. The technology side has seen significant advances with the introduction of structured programming, structured design, formal verification, automatic code generators, diagnostic compilers and libraries of reusable software modules [Ref. 2:p. 1]. The managerial side, in comparison, has not seen the same level of significant progress made. Although some managerial advances have been made in computer-aided software engineering, estimation tools, software metrics and quantitative models of software project dynamics, these fields are either still too new or too rigidly calibrated to a particular organization to be generally useful throughout the software industry.

The Systems Dynamic Model of Software Project Management (SDM) is one of the new quantitative models of software project dynamics that is attempting to gain some valuable insight into the managerial side of developing software systems [Ref. 2:pp. 4-9]. It is a comprehensive simulation model of the software development process that integrates both the management type functions (e.g., planning, controlling and staffing) with the software production type activities (e.g., design, coding, reviewing and testing). [Ref. 2:pp. 6-9]

The simulation model is a viable laboratory tool that can be used for controlling experimentation in the software project management field to test various management hypotheses. Conducting this experimentation without the use of a simulation model has proven to be too costly and time consuming. Furthermore, the isolation of the treatment and the analysis of the results for a large complex software project can be exceedingly difficult. The use of a simulation model permits less costly, less time consuming and perfectly controlled experimentation possible. [Ref. 3:p. 10] Indeed:

> The effects of different assumptions and environmental factors can be tested. In the model system, unlike the real systems, the effect of changing one factor can be observed while all other factors are held unchanged. Such experimentation will yield new insights into the characteristics of the system that the model represents. By using a model of a complex system, more can be learned about internal interactions than would ever be possible through manipulation of the real system. Internally, the model provides complete control of the system's organizational structure, its policies, and its sensitivities to various events. [Ref. 4:p. 1]

The valuable insight provided by the SDM spans four managerial issues. First it can be used as an aid in understanding the software development process through the manager's ability to track, store and plot large amounts of project data, quickly and efficiently. The manager's ability to replay the simulation with a change in a single variable promotes a more comprehensive understanding of the interrrelationships of the software development variables. Once calibrated to an organization, the model can also be

used as an aid in the actual management process. By selectively changing variables in the model to reflect possible upcoming changes in the organization's software development process, the manager can determine the effects of the change on the schedule and cost of a project before the change gets implemented.

The use of the SDM gaming interface provides the last two major uses for gaining valuable insight into the software management process. The gaming interface can be used as a training tool for inexperienced software project managers. The gaming interface allows the trainee to halt the simulation at specified time intervals and make changes to the software development variables. This interaction enables the trainee to see the immediate impact of his managerial decisions. The final managerial insight involves using the gaming interface to conduct experiments on how software project managers make project management decisions during the development process. A number of software project managers can run the exact same project through the gaming interface. Their results can be compared to each other to investigate an endless list of software project management concerns and theories. A major source of the concerns in software development today border on the heuristics and biases that go into the software project manager's decision making processes. Software project management, after all, involves decision making under great

uncertainty. As Schlender added in his final comments, "software remains the most complex and abstract activity man has yet contrived." [Ref. 1:p. 112]

B. PURPOSE OF RESEARCH

The objective of this thesis is to design, construct and execute three experiments, using an enhanced version of the SDM gaming interface, to investigate software project manager heuristics and biases. Each experiment will address a specific software project manager heuristic or bias that can prevent a software project from being reliably completed with the best mix of effort expended and project duration.

"Anchoring" is the first heuristic investigated. "Anchoring" is a heuristic in which people unduly rely on a given variable's initial estimate when making future adjustments to the variable. "Anchoring" reduces the complex tasks of assessing probabilities and predicting future estimates in an uncertain world by enabling an individual to use much simpler judgmental operations. The use of an "anchor," though, can sometimes lead to severe and systematic errors [Ref. 5:pp. 35-38]. Specifically, the experiment will investigate whether or not software project managers "anchor" revised estimates of overall staff productivity towards a given initial estimate.

The second experiment looks at how an incorrect initial estimate of needed effort affects the manner in which a software project manager makes staffing decisions during the

development phase of a project. This experiment investigates the software project manager's bias towards fulfilling the prophecy of the initial estimate.

The final experiment explores the "Social Loafing" phenomenon. As applied to software project management, the experiment compares the performance of software project managers that assume control of a project at its inception with those that assume control at some point into the project lifecycle. The comparison is made through an analysis of the final effort expended and duration of a project achieved through the software project manager's use of the available work force.

C. SCOPE OF RESEARCH

The scope of this research includes the design, construction, preparation of documentation and software, execution and analysis of the software project manager heuristic and bias experiments. The design consisted of identifying the dependent and independent variables that, when controlled in an experiment, will best achieve the desired objective. The construction phase consisted of tailoring the SDM to emulate a specific project and organization under the guidelines of the controlling experimental variables. The gaming interface was enhanced for each experiment to improve the display of reports, provide better control of the experiments execution path and

to specify important directions to the experimental subjects.

The first part of the preparation phase entailed writing a documentation package for all the groups in each experiment. Then the software for each experiment was compiled and downloaded to floppy disks for each subject and then added to the documentation package. Each subject had his own package that included the documentation and software for each of the three experiments. The execution phase was conducted over two days and in two locations each day due to the limited number of microcomputer resources available. The analysis phase consisted of evaluating the experimental data with the SAS statistical system.

D. ASSUMPTIONS

The subjects in these experiments were fifth and sixth quarter graduate students studying in the computer systems management and computer science curriculums at the Naval Postgraduate School. Through the use of a pre-experiment questionnaire, Appendix A, it was determined that none of the students had any extensive experience in project management or software development. Even though the subjects are not active software project managers, the results of the experiment and the conclusions made from them are assumed to parallel those that would be found in the software industry. This assumption is given validity by the work of William Remus [Ref. 6:pp. 19-25]. His study on

using graduate students as surrogates for similarly educated managers in experiments on business decision making found that there were no significant differences between graduate students and business managers in making production scheduling decisions. Although software project management decisions are somewhat different from production scheduling decisions, they are similar enough to apply his findings to the assumption that graduate students are acceptable surrogates in this thesis's experimental investigation.

The students were not monetarily compensated for their participation in this experiment, in violation of accepted experimental microeconomics protocols [Ref. 7:p. 24]. They were told, however, that their quality of participation accounted for ten percent of their grade in the Software Engineering Management course they were concurrently taking.

E. THESIS ORGANIZATION

Chapter II is an in-depth review and analysis of the "Anchoring" experiment. Chapter III analyzes the experiment that examines the effects of an incorrect initial estimate of effort needed on a software project manager's staffing decisions during the project's development. Chapter IV describes the "Social Loafing" experiment and analyzes the experimental results. Chapter V summarizes the significant conclusions presented in Chapters II-IV and provides lessons learned and future direction for follow-on theses.

II. INVESTIGATION OF "ANCHORING" IN SOFTWARE PRODUCTIVITY ESTIMATION

A. IMPORTANCE OF THE "ANCHORING" PHENOMENON IN SOFTWARE PROJECT MANAGEMENT

A major portion of a software project manager's job revolves around being able to estimate future events. The uncertainty of these events (e.g., personnel turnover, requirements changes, anticipated needed staffing level, complexity, staff productivity, project duration, cost, etc.) and the inability of the software project manager to predict all these events accurately make developing software an extremely risky venture. Farquhar explains the significance of poor estimation on the software development process:

> Unable to estimate accurately, the manager can know with certainty neither what resources to commit to an effort nor, in retrospect, how well these resources were used. The lack of a firm foundation for these two judgments can reduce programming management to a random process in that positive control is next to impossible. This situation often results in the budget overruns and schedule slippages that are all too common. [Ref. 8:p. 1]

In addition to the uncertainty of future events a number of contributing factors degrade the estimation process. Until built, software is an abstract entity. There is no blueprint for success that can show all its parts. Software is becoming more complex and is frequently attempting to break new ground. There is a severe lack of estimation experience in software project managers. The few cost and

schedule estimation tools available must be calibrated to a frequently changing organization in order to be useful.

Tversky and Kahneman have noted that when faced with the outcome of predicting complex and uncertain events, "people rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations." [Ref. 5:p. 34] "Anchoring" is one of the heuristic principles that is very important in the software development process. "Anchoring" is a bias in which future adjustments to a variable are unduly influenced by an initial or earlier value. Giving people different starting values, or "anchors," for the exact same problem yields different future estimates based on the given "anchor" [Ref. 5:p. 38].

Given the widely-documented problems in predicting and using initial estimates in software development, "anchoring" to these initial estimates during the project lifecycle can be disastrous! The project data generated during the lifecycle process can provide keen insight to what is actually occurring during the project's development. The problem with the data is that they are large, difficult to collect and time-consuming to analyze. Using these data to revise estimates would obviously provide better results than relying on the "anchor."

B. EXPERIMENTAL OBJECTIVE

This experiment investigates whether or not software project managers "anchor" their estimates of overall staff productivity towards a given initial estimate. Overall staff productivity is defined as tasks completed per man day.

C. EXPERIMENTAL DESIGN

1. Basic Framework

The basic framework of this experiment was set up to be similar in many ways to the flight simulators that pilots use to mimic flying an aircraft from takeoff at point A to landing at point B. Instead of flying an aircraft, though, this simulator mimics the life of a real software project from the start of the design phase until the end of testing. In less than an hour the subjects would live through a project's lifecycle. The subjects would be more than outside observers.

Their role was defined not to be that of a project manager, but rather they played the role of a valuable assistant to the manager (i.e., using the flight simulator analogy again, their role was that of the flight engineer).

Specifically, their role involved tracking the project's progress using a number of reports that were produced for them by the model at different intervals during the project, and they were required to make their best estimate of the project team's overall average productivity

(in Tasks/man-day). They were told their estimate would be critically important to the project manager, since he/she would use this information to make the necessary adjustments to the project's staff and schedule. In reality, the model was designed such that the subject's estimates of the overall average productivity had no influence upon the project's actual development. The reason for this was to ensure that the model provided identical behavior for each subject. Identical behavior for each subject was necessary in order to test for the presence of "anchoring" between and within the experimental groups. The subjects, on the other hand, had to feel like they were performing a meaningful assessment of the work being completed. Telling them that their assessment would be used by the "simulated project manager" to finish the project in the most economical and efficient manner was the only way to ensure that they tried their best in accomplishing the task at hand.

Overall staff productivity was chosen as the dependent variable due to its relative independence from the other managerial decisions made during project development. By relative independence, I mean that I was able to sufficiently hide the model's disregard of their productivity estimate so that the subjects would not detect that their input was not being used by the model. To aid in this deception, the number of estimate revisions solicited from the subjects was held to four; one after the completion

of the Design phase (100 work days into development), one after each of two coding and testing increments (200 and 300 work days into development) and the last at the completion of testing the third and final increment (385 work days). The subjects believed that their input was one of several factors taken into consideration by the model in determining the work force level needed to complete the project within the schedule constraints. With the solicitations for revised productivity estimates coming every 100 calendar work days (five months), the subjects would not be able to determine if their 100 day-old estimate had any influence on the project's current staffing level.

The software project used in this experiment was a real software project developed at NASA in the early 1980's. It contained 610 tasks and took 2064 man days of effort to complete. The actual overall staff productivity was approximately 0.30 tasks per man day.

2. Experimental Groups

The subjects were randomly divided into three experimental groups of 12 subjects each. The randomness was accomplished through assigning a two digit value from a random number table to each subject on an alphabetical class listing. A random number from one to 33 placed the subject in one group, 34 to 66 in another group and 67 to 99 in the third group. The number zero was discarded. Once a group attained 12 subjects, its corresponding random numbers were

also discarded. The control group received an initial productivity estimate of 0.30 tasks/man-day. The other groups were given initial estimates higher and lower than the perfect estimate. The under-estimated group was given an estimate that was 33% below the actual overall productivity rate, namely 0.20 tasks/man-day. The actual productivity rate of 0.30 tasks/man-day was 33% below the high group's estimate of 0.45 tasks/man-day.

3. Documentation

The documentation given to each group was exactly the same except for the page entitled "Management's Initial Project Estimates." This page was altered for each group to reflect the difference in the initial estimate of overall staff productivity. Appendix B contains a copy of the documentation package with each "Management's Initial Project Estimates" page attached.

4. Dynex Gaming Interface Control File

The actual SDM Model used in the experiment was identical for each group. The addition of dummy variables, to reflect the given initial estimates, and minor changes to the enhanced Dynex gaming interface control file created the illusion that each group was working on a different project.

The Dynex gaming interface control file was enhanced to include an initial screen of instructions before continuing with the experiment. In addition, the control file was altered to solicit only the revised estimate of the

staff's overall average productivity. The gaming interface output was changed from standard plots to the report screen shown in Figure 2-1. Each subject saw the same report screens and values at each stoppage of the experiment no matter which group they were in.

CURRENT INTERVAL STATISTICS: Elapsed Time = 40

INITIAL ESTIMATES: (These will not change throughout the project)

Project Size	500		Tasks
Man-day Cost	2330.00		Man Days
Project Duration	345		Days
REPORTED STATISTICS at time =>	40		Days
% Project Reported Complete	8.43		Percent
Updated Size of Project	500		Tasks
Total Number--Fulltime Equiv Staff	6.5		Fulltime
Effort Expenditures to Date:			
Development Activities	215.98		Man Days
Design and Coding		154.61	Man Days
Rework (i.e., fixing errors)		28.97	Man Days
Quality Assurance		32.40	Man Days
Testing	0.00		Man Days
Total Man Days Expended	215.98		Man Days
New Est of Duration (start--end)	345		Days
Max Tolerable Project Duration	400		Days

Write your new desired staffing level on the documentation sheet provided and press <ENTER>

Figure 2-1 Sample Project Status Report

5. Experiment Execution

The "Anchoring" experiment was executed first, followed by Experiment two and finally the "Social Loafing"

experiment. The subjects were initially gathered in a classroom and presented with the documentation for the "Anchoring" experiment. After reading the documentation package, they were given a 20 minute presentation that included a definition of productivity, an insight into some of the considerations that should go into the revised productivity estimate since there is no clear-cut calculation that will yield the correct answer until the final project statistics are known, a reminder that early reported project statistics generally follow the budgeted and not the actual progress, a warning to work alone and instructions on how to play the game. Following the presentation, the subjects were given a brief on-line view of how the gaming interface worked. This enabled them to clearly see how to work the experiment and offered them the opportunity to ask any pre-experiment questions.

Due to the limited microcomputer resources available, the subjects were sent to one of two labs depending on which "Social Loafing" experimental group they were in. A special seating arrangement was used in each lab to minimize the interaction between "anchoring" groups.

Each subject was required to determine a revised estimate of the staff's overall average productivity at each stoppage of the simulation. The revised estimate had to be entered into the model through a solicitation screen and written on a special estimation sheet that was submitted to

a lab attendant before the subject was allowed to continue the simulation. The submission of the written estimate at the time of the simulation stoppage is important because, as the project finishes, the subject can easily calculate the actual overall average productivity from the final project statistics. Collecting the written estimates during the experiment prevents a subject from changing previous estimates. Upon completion of the project, each subject was required to briefly specify the method they used to calculate their revised estimates.

D. "ANCHORING" EXPERIMENT RESULTS AND ANALYSIS

The raw results from the experiment contain revised estimates of the staff's overall average productivity for 34 subjects. Six observations were excluded from the final analysis. These were excluded due to the subjects admitted misunderstanding of what was required from them during the experiment. Each of the six subjects had observations that significantly deviated from their group's mean responses. Appendix C contains a list of the students assigned to each group and reasons, if any, for their observations being excluded from the final analysis. Table 2-1 lists the subject's productivity estimates made during the experiment that were used in the final analysis.

Figure 2-2 is a plot of the three groups mean estimates of the staff's overall average productivity from the initial estimate up to and including the third revised estimate.

TABLE 2-1

"ANCHORING RESULTS USED IN FINAL ANALYSIS

NAME	INITIAL ESTIMATE	TIME 100	TIME 200	TIME 300	PROJECT COMPLETION
Acton	.2	.12	.16	.15	.29
Ellis	.2	.279	.338	.323	.302
Johnson	.2	.33	.37	.36	.29
Peterson	.2	.225	.15	.15	.15
Rouska	.2	.35	.25	.1	.2955
Shuman	.2	.3301	.3752	.3584	.2921
Sweitzer	.2	.15	.18	.25	.30
Taylor	.2	.33	.2	.23	.295
Zeiders	.2	.27	.28	.23	.19
Beedenbender	.45	.5	.37	.33	.29
Bell	.45	.4	.5	.51	.296
Bischoff	.45	.52	.5	.45	.3
Clemens	.45	.51	.44	.4	.3
Drummond	.45	.33	.18	.3	.29
Garrabrants	.45	.33	.4	.38	.3
Mostov	.45	.4	.45	.4	.4
Myers	.45	.55	.4	.45	.2955
Rassatt	.45	.37	.35	.32	.27
Sablan	.45	.35	.3	.25	.27
Ash	.3	.26	.35	.28	.3009
Banham	.3	.29	.29	.3	.296
Chase	.3	.33	.35	.3	.29
Lekey	.3	.27	.3	.31	.29
Newton	.3	.35	.41	.4	.292
Sawyer	.3	.318	.405	.459	.295
Schwind	.3	.33	.38	.36	.2955
Spaulding	.3	.33	.375	.358	.292
Triebwasser	.3	.32	.337	.441	.343

The final estimate of the overall average productivity was a straight calculation and provides no insight into the "anchoring" phenomenon. It is not included in the plot or accompanying analysis.

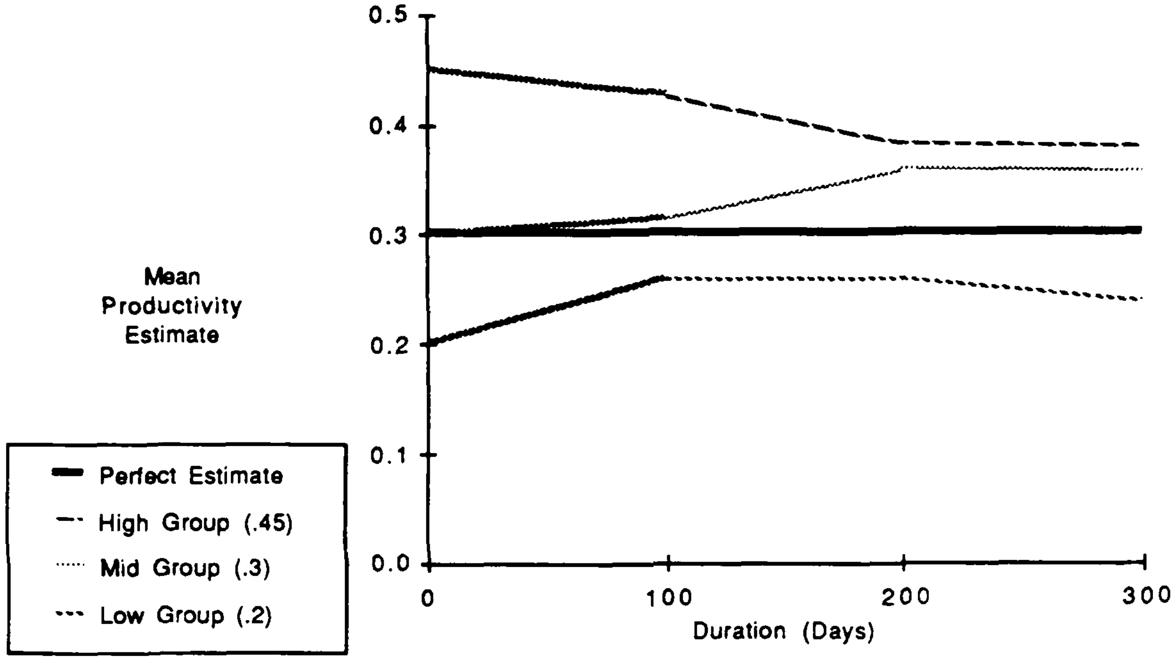

Figure 2-2 Groups Revised Productivity Estimates

A repeated measures analysis of variance test was used to examine the experiment results. The SAS control file used to analyze the data is listed in Appendix D. Table 2-2 lists the results from multivariate repeated measures tests.

TABLE 2-2

RESULTS OF REPEATED MEASURES TESTS

Test	F-Value	Prob > F
No time effect	F(2,24) = 0.2	0.8161
No time and group effect	F(4,48) = 2.0	0.1052
Between subjects effects	F(2,25) = 11.4	0.0003

The first test determines the effect of time on the groups revised estimates. The null hypothesis is that there is no time effect on the subjects revised estimates. In other words, the lines connecting the groups mean estimates from time 100 to time 300 are horizontal. The high p-value of 0.8161 clearly prevents the rejection of the null hypothesis. Referring back to Figure 2-2, this test states we cannot say that the lines connecting the groups estimates are significantly non-horizontal. The groups estimates do not change significantly over time alone. [Refs. 9:p. 190; 10:pp. 478-483]

The next repeated measures test was a multivariate test to determine the level of no time and group effect. Referring back to Figure 2-2 again, this can be interpreted as saying the lines for the three groups after time 100 are parallel to each other. The p-value of 0.1052 is above the desired significance level of 0.05, therefore the null hypothesis cannot be rejected.[1] The lines are not significantly non-parallel and, therefore, the individual groups did not raise or lower their productivity estimates

---

[1]A univariate test of the same measure yielded a p-value of 0.0366 which meets the desired significance level of 0.05. The rejection of the null hypothesis would signify that the groups did abandon their "anchor" over time. A sphericity test to determine the worth of the univariate test resulted in a p-value of 0.0008. The dramatic rejection of the sphericity test casts much suspicion on the validity of this univariate result. [Ref. 10:pp. 605-606]

over time any differently than the other groups in the experiment.

The between subjects effects test, with a null hypothesis that the groups mean revised estimates over time are the same, yielded a p-value of 0.0003. The result of this test is a dramatic rejection of the null hypothesis. The different mean estimates calculated for each group are significantly different. Again referring to Figure 2-2, the null hypothesis states that the three lines depicting the mean productivity estimates over time for each group are not significantly different from each other. The rejection of the null hypothesis demonstrates that the lines are significantly different and that each group's mean estimates were different from those of the other groups.

E. CONCLUSIONS

Although the groups with the low and high initial estimates approached the correct estimate of 0.30 tasks/man-day, the results of the between subjects test, Table 2-2, and the plot of the groups mean productivity estimates, Figure 2-2, state that the groups did "anchor" their revised estimates towards the given "anchor." The multivariate test of no time and group effect shows that the groups did not abandon their "anchor" over time. This result is somewhat surprising. I fully expected the subjects to approach the perfect estimate of 0.30 tasks/man-day as the project neared completion at time 300. Their reluctance to significantly

change their revised estimates, even when presented with nearly completed project data, proves that software project managers rely on available heuristics to reduce the complexity of decision making under uncertainty. In this experiment, the software project managers "anchored" to a given initial estimate of overall average productivity to reduce the complex task of determining a revised estimate of the overall average productivity to a simpler judgmental operation.

III. EXPERIMENT TWO: THE EFFECT OF DIFFERENT INITIAL COST ESTIMATES ON STAFFING DECISIONS

A. DIFFERENT INITIAL ESTIMATES CREATE DIFFERENT PROJECTS

Research findings and experimentation using the SDM indicate that staffing decision are significantly influenced by the pressures and perceptions that project schedules produce. Figure 3-1 is a causal loop diagram that shows how project estimates directly influence the hiring and firing decisions throughout a project's development phase [Ref. 3:p. 12]. Project estimates, along with the progress made on the project, directly affect the work force hiring and firing decisions. If the estimates and progress made dictate the need for a larger work force, this decision will lead to increased communication and training overheads on the project. This will, in turn, decrease the staff's productivity. The reduced productivity then affects the progress level that will be achieved and, in turn, lengthens the revised project estimates. The initial project estimates therefore have a strong influence on hiring and firing decisions, productivity, and communication and training overheads [Ref. 3:p. 13].

The project used in the SDM experimentation into project manager staffing decisions was the real life DE-A project developed by NASA. The DE-A was one of the original projects used to validate the SDM model. During the

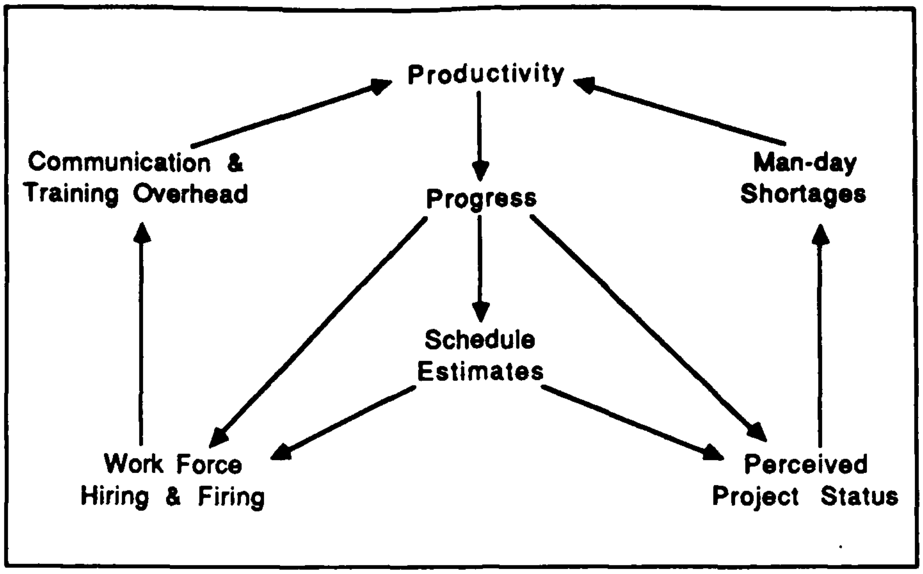

Figure 3-1 Causal Loop Diagram

validation process, the SDM simulation, using the initial estimates developed by NASA, closely mirrored the actual project variables history. [Ref. 3:pp. 8-10]

The SDM experimentation involved running the DE-A project through the model twice. Each run was made under the exact same conditions except for the initial project cost estimate. The initial project estimates given to the model were generated using two different estimation tools; WHIZ and COCOMO. Table 3-1 is a summary of the initial estimates and final project results for the two model runs and for the actual NASA-developed DE-A project. [Ref. 3:p. 12]

Clearly evident in Table 3-1 is the wide range between the project cost estimates for the two estimation tools.

TABLE 3-1

ESTIMATES AND FINAL RESULTS FOR DE-A PROJECT

	DURATION (Days) Estimate	DURATION (Days) Final	COST (Man-days) Estimate	COST (Man-days) Final
WHIZ	237	243	3500	3516
COCOMO	237	316	1305	2588
ACTUAL (NASA)	320	380	1100	2200

Neither comes particularly close to the actual cost of 2200 man-days. In fact both have a relative error of over 40%. Both of these tools performed miserably if you subscribe to the notion that the actual project totals are independent of the initial estimates. The SDM simulation runs using the WHIZ and COCOMO initial estimates, though, challenge the notion of independence. The fact that the initial duration estimates were identical enabled the experiment to focus on how the different initial man-day cost estimates affected the final project statistics. The final project results for the model runs show how the different project cost estimates do indeed create projects with different final costs and durations.

In addition to creating different final project totals, the different initial project estimates had a profound effect on the work force level used during the development phase. Figure 3-2 shows the work force used over time by the model for each set of initial estimates. Due to

COCOMO's under-estimation of the man-days needed, its curve has a dramatic rise toward the end of the development phase when management realizes that they still have a significant number of tasks left to complete. The WHIZ curve meanwhile shows an early build-up of personnel that remains fairly stable throughout the development phase. [Ref. 3:p. 14]

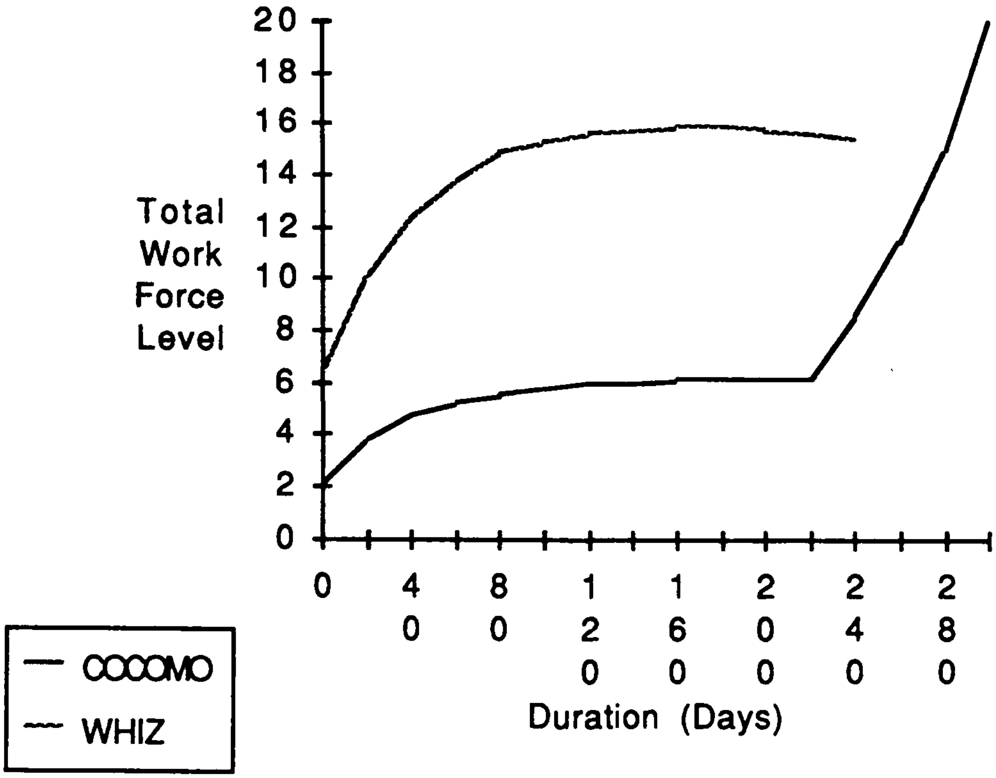

Figure 3-2 WHIZ & COCOMO Work Force Curves

The SDM model mirrored the actual results based on NASA's original under-estimated project cost and duration. Would the model's staffing decisions have mirrored the DE-A

project if NASA's original project cost and duration were initially over-estimated? This question is important because it forms the basis for studies that are currently looking into the use of historical project data for schedule and cost estimation.

To answer the question, and verify the model's staffing decisions when faced with over- or under-estimated initial project costs and durations, real software managers must be allowed to make staffing decisions for projects that are identical except for the initial man-day cost. Developing the DE-A project again with different managers and different initial estimates is too costly and unrealistic. The SDM gaming interface, though, provides a logical and suitable alternative.

B. EXPERIMENTAL OBJECTIVE

The objective of this experiment is to compare the desired staffing level decisions, throughout the development phase, of software project managers managing identical projects whose only difference is that their man-day cost is initially under-estimated, over-estimated or perfectly estimated.

C. EXPERIMENTAL DESIGN

1. Basic Framework

The basic framework of this experiment was to create identical SDM project scenarios that differ only in their

initial man-day cost estimates, and to track the staffing decisions of software project managers throughout the project's development phase. The DE-A project from NASA was used due to its availability and use in the previous SDM experimentation explained above.

Unlike the "anchoring" experiment in which the subjects played the role of "flight engineer" and provided estimates of the overall average staff productivity, in this experiment they have been promoted to "Captain," also known as project manager, and were required to make the project's staffing decisions. The subject's task was to use the reports, on resources used to date, work accomplished, current staffing level and elapsed time, generated by the model at different points during the development phase to determine a desired staffing level for the remainder of the project that they felt provided the best compromise between finishing on an acceptable schedule while avoiding an excessive cost overrun.

The only project management decision solicited from the subject during the experiment was for the desired staffing level, also known as work force level sought (WFS). WFS is the staffing level that the project manager desires. As in a real project management situation, the model does not give the project manager absolute control over the work force level. Turnover, promotions, work force ceilings,

transfers, hiring and assimilation delays prevent the manager from always getting the <u>exact</u> work force he wants.

Determining the perfect estimate for the DE-A project required running the actual DE-A project results through a normalization engine to obtain normalized initial estimates. Abdel-Hamid provides an in-depth look at how this procedure can be applied to the DE-A project results in [Ref. 3:pp. 16-19]. He found that the perfect cost estimate given a desired project duration of 380 days is 1900 man-days. The 380 day project duration was the actual duration of the DE-A project.

For this experiment the initial project duration was set to 380 days with an acceptable completion range of only 370-390 days. The maximum tolerable completion date was also limited to just 390 days. It was necessary to tighten the completion range so that a more reasonable comparison of the desired staffing level decisions for the various groups could be made.

Finding the perfect estimate with the normalization engine assumes that the project's size (in DSI) be correctly estimated at the project's initialization. The SDM's estimated project size throughout the development phase, therefore, is the actual size of the DE-A project, 24,400 DSI. All the subjects were given the same actual final size in all of the current project statistics reports.

2. Experimental Groups

Although this experiment was conducted prior to the "social loafing" experiment described in Chapter IV, the design of the experiment and the assignment to experimental groups occurred after the "social loafing" experiment was finalized. The subjects assigned to these experimental groups, therefore, were randomly selected from the two groups in the "social loafing" experiment. Originally there were only three experimental groups in this experiment. For each 18-person "social loafing" group, six subjects were randomly sent to each of the three groups using a random number table. Just prior to the execution of the experiment, another group was added. Three subjects from each group were then randomly selected to be part of the fourth group.

The groups were designated "G-number" with the number corresponding to the number of man-days in the group's initial estimate of project cost. The perfectly estimated group was designated "G-1900" for an initial estimate of 1900 man-days. There were two groups that received over-estimated initial project costs. One group was given an estimate of 2470 man-days, "G-2470," whereas the other group was given an estimate of 2185 man-days, "G-2185." The under-estimated group was "G-1460." Appendix E contains a list of the subject assignments to the four experimental groups.

3. Documentation

With the exception of the given initial estimate for the man-day cost, the documentation for each group was identical. Appendix F contains a sample documentation package and the initial estimate sheets for all four groups.

4. Dynex Gaming Interface Control File

The DE-A project was used in the SDM for this experiment. The model's project variables were identical for each group except for the initial man-day cost. The initial man-day cost was set to match the subject's particular experimental group.

The gaming interface control file was the same for each group. Initially it showed a page of instructions, as listed in Appendix G, for running the experiment. Then it solicited the subject for his initial desired staffing decision. After simulating the project for 20 days, a current project statistics report was displayed (see Appendix H). After deciding on a new desired staffing level, the subject would enter it into the model and simulation would continue in 20-day increments until the project was completed.

5. Experiment Execution

After completion of the "anchoring" experiment, the subjects were given the documentation package for this experiment. Any questions concerning the experiment were answered prior to allowing the subjects to boot the gaming

interface control file. The subjects were allowed to work at their own pace. The only requirement was that they make a desired staffing level decision at each 20-day period. The decision also had to be entered at the simulation prompt and written on the experiment documentation sheet (Appendix I). The documentation sheet allowed the subjects to check their progress over time and aided in the analysis of the results.

D. EXPERIMENT "TWO" RESULTS AND ANALYSIS

The results for experiment two consist of the desired staffing level decisions and the final man-day cost and project duration values for eight subjects in groups "G-2185" and "G-2470" and nine subjects in groups "G-1460" and "G-1900." The SAS control file used to create the statistics analyzed in this section is listed in Appendix J.

Figure 3-3 is a graph of each group's desired staffing level decisions. The plot extends from the initial choice at time zero until the time period immediately following the group's mean final project duration value. For example, in group "G-1900" the subjects' durations ranged from 290 to 380 days with a mean for the group of 346 days. The plot for group "G-1900" ends at the time period following 346, namely 360 days. Stragglers that finished after time period 360 are not plotted due to the relative distortion their small sample size inflicts on the graph. Each group's initial desired staffing level decision at time zero was

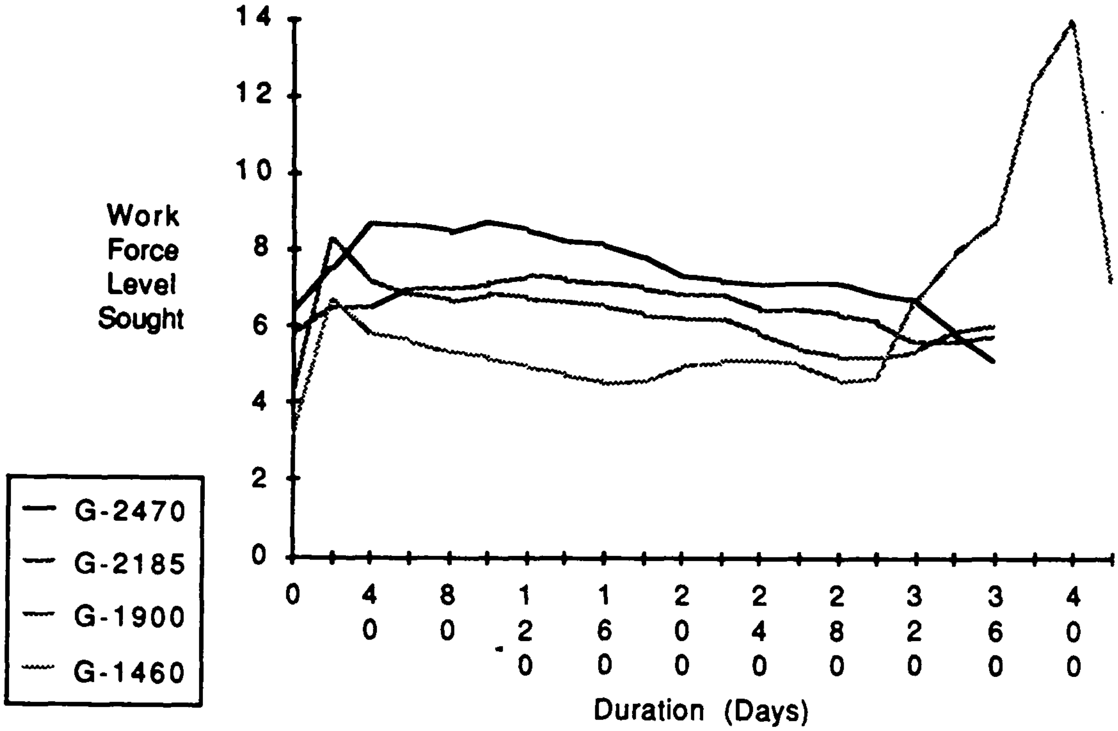

Figure 3-3 Comparison of Groups WFS Decisions

followed by a significantly larger desired staffing level at time 20. The reason for the low initial values is that the subjects were given a core team of one and one-half full-time equivalent experienced workers that could be used at the beginning of the project. This core team was not large enough to complete the project on schedule, but a significant number of subjects used that number as an "anchor" for their initial desired staffing decision at time zero instead of calculating the work force level that they really needed. Upon seeing what the low initial desired staffing level figure did to their estimated project duration on the

current project statistics report at time 20, most of the subjects significantly raised their desired work force level. For the groups with the lowest estimates, "G-1460" and "G-1900," they increased the work force level to a degree that the time 40 project report had them completing the project too early. The probable cause for this over-correction was the subject's failure to calculate an assumed productivity for the work force. Upon seeing the severe schedule problem created, on the average a 50% increase in duration, they innocently just doubled their desired staffing level. From time 60 onward, the groups settled into a stable pattern. The under-estimated man-day cost given for "G-1460" forced the subjects to dramatically raise their WFS levels near the end of the development phase when they realized that they still had much coding and all the testing left to complete.

The final project durations for the groups provide an expected result, namely that a project developed using an under-estimated initial man-day cost will take significantly longer to complete than a project that was accurately or over-estimated. Table 3-2 is a nonparametric analysis of the final project duration of the under-estimated group, "G-1460," compared to the final project durations of the combined perfect and over-estimated groups, "G-1900," "G-2185" and "G-2470." A formal normality test (in SAS it is the normal option under procedure univariate) of the final

project duration values rejected the assumption of normality and necessitated the use of the nonparametric Wilcoxon Rank Sum Test. [Ref. 9:pp. 117-118]

TABLE 3-2

NONPARAMETRIC ANALYSIS OF PROJECT DURATION

Group	Mean Project Duration	N	Wilcoxon Scores Sum	Wilcoxon Scores Mean
"G-1460"	402	9	240	26.67
"G-1900, 2185, 2470"	348	25	355	14.20

Wilcoxon 2-Sample Test S = 240 Z = 3.2083
Prob > Z = 0.0013
Kruskal-Wallis Test CHISQ = 10.42 DF = 1
Prob > CHISQ = 0.0012

The combination of groups in this particular analysis was important because it was the only way to isolate the group that was managing the project based on an under-estimated project cost. The null hypothesis was that the mean project duration for subjects that received an under-estimated cost was equal to the mean project duration of the subjects that did not receive an under-estimated cost. The subjects in the three groups "G-1900," "G-2185" and "G-2470" fall into the category of not receiving an under-estimated cost. Although the under-estimated group only had nine subjects compared to the 25 in the not under-estimated group, the Wilcoxon Rank Sum Test does not require groups of equal sizes. [Ref. 9:p. 196]

The p-value of 0.0013 significantly rejects the null hypothesis that the mean project durations are equal. The mean project duration for the under-estimation group "G-1460" is significantly higher than the combined mean project duration for the other non under-estimated groups.

A repeated measures analysis of the WFS decisions was made for decisions from the initial WFS decision at time zero until time 340. The repeated measures analysis ended at time 340 to prevent the loss of too many observations due to missing values. A subject could not be included in the repeated measures analysis if he finished prior to time 340 due to the non-availability of a WFS decision at time 340.

Table 3-3 lists the results of the repeated measures tests. The first test has the null hypothesis of no time effect on the WFS decisions. The p-value of 0.3828 prevents the rejection of the null hypothesis. Referring to Figure 3-3, the lines are not significantly non-horizontal. The dramatic rise in the WFS line of group "G-1460" comes after the termination point for the repeated measures test.

TABLE 3-3

RESULTS OF REPEATED MEASURES TESTS

Test	F-value	Prob > F
No time effect	$F(17,5) = 1.4$	0.3828
No time and group effect	$F(51,16) = 1.0$	0.5264
Between subjects effects	$F(3,21) = 63.8$	0.0001

The result of the test for no time and group effect was a p-value of 0.5264 which again prevents the rejection of the null hypothesis. The lines in Figure 3-3 are not significantly non-parallel from time zero through time 340. After time 340 it is clear from Figure 3-3 that group "G-1460" takes a significant upward turn that is not evident in any of the other groups.

The final repeated measures test shows the between subjects effect. The p-value of 0.0001 significantly rejects the null hypothesis. The individual group lines in Figure 3-3 are significantly different from each other from time zero through time 340.

In addition to analyzing how the groups compared to each other, the group WFS decisions were compared to how the SDM determined the WFS under the exact same conditions as each group. Table 3-4 and Figure 3-4 depict how the groups mean project cost and duration compare to the SDM values.

TABLE 3-4

GROUPS FINAL COST AND DURATION

Group	Group Cost	SDM Cost	Group Duration	SDM Duration
G-1460	2031	2016	402	420
G-1900	1964	1876	346	380
G-2185	2104	2116	352	375
G-2470	2346	2268	346	365

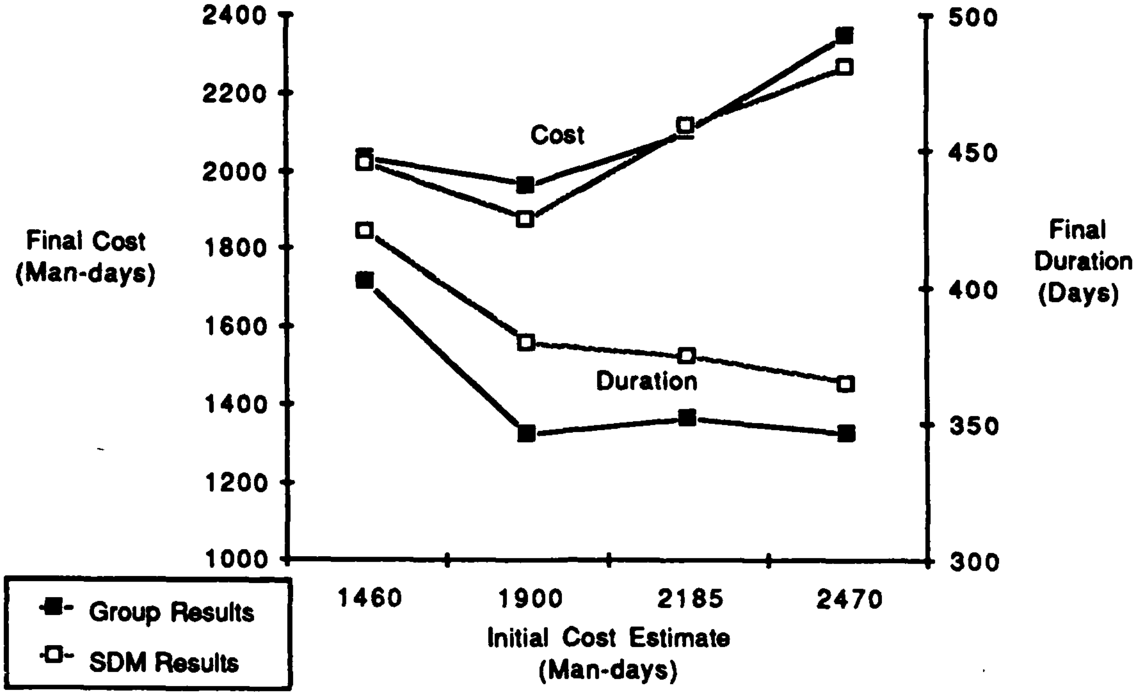

Figure 3-4 Group vs. SDM Final Project Value Comparisons

The final project duration for the model's run under the same conditions as group "G-1460" is much higher than the other three groups. This finding is consistent with the one observed when the groups ran the experiment. Under-estimation leads to a longer project duration.

Figure 3-5 is a graph of the WFS decisions for the SDM runs for each of the four initial estimates used by the experimental groups. This graph compares favorably with Figure 3-3, the graph of the groups WFS decisions, except for the groups instability in the initial three time periods. Although there is no statistical test to prove the significance of the comparison, it seems that the higher the

initial estimate of man-day cost the higher WFS decisions over time for both Figure 3-3 and Figure 3-5.

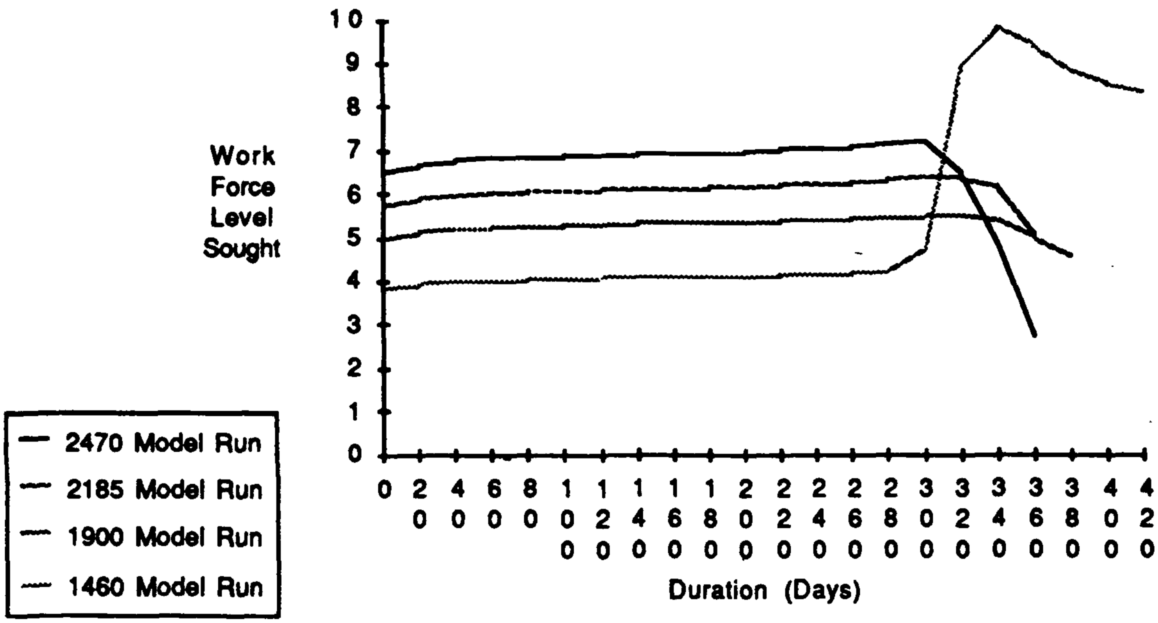

Figure 3-5 Comparison of SDM WFS Decisions

The 1460 model run has a dramatic rise near the end of the development phase in similarity with the group G-1460's trend. Figure 3-6 depicts the closeness of the fit between the group's and model's response. Similar plots of the other three groups, Figures 3-7 through 3-9, yield the same results. In all cases the subjects jumped out to a larger WFS decision in the early stages of the development phase then gradually approached the model. A comparison of the plots does not show any significant differences between the

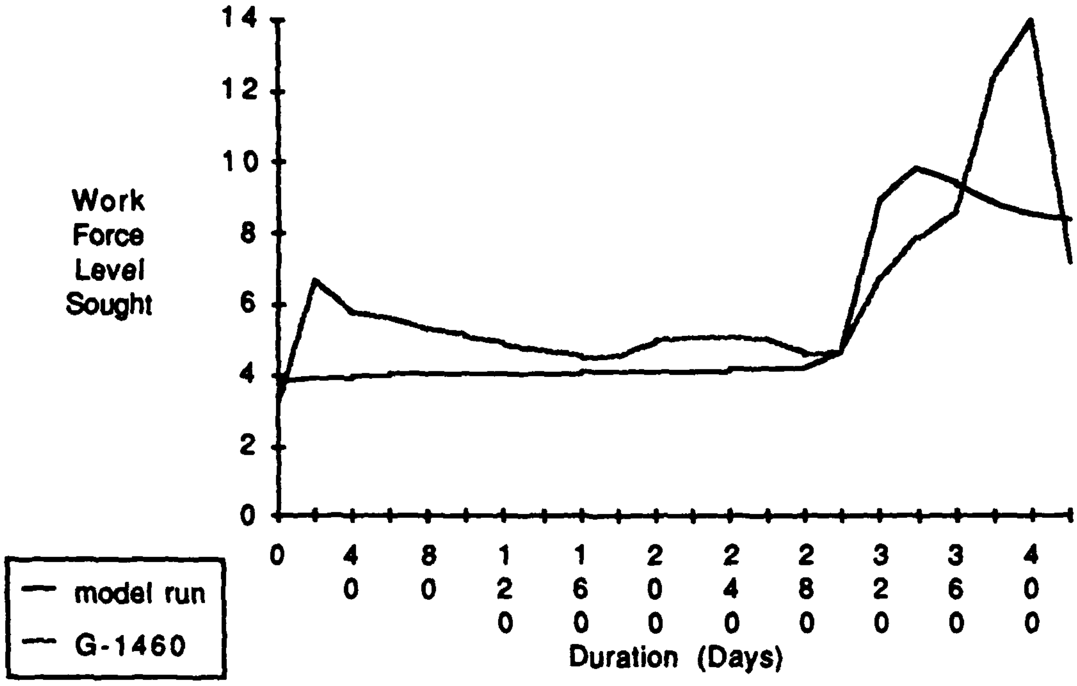

Figure 3-6 "G-1460" WFS Decisions vs. SDM

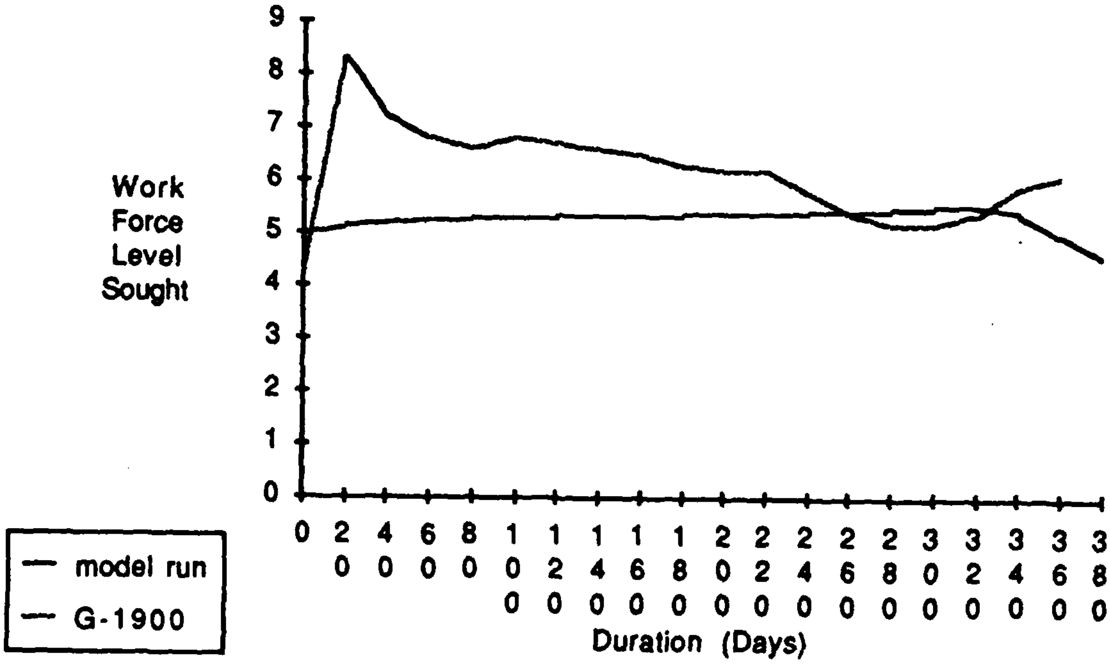

Figure 3-7 "G-1900" WFS Decisions vs. SDM

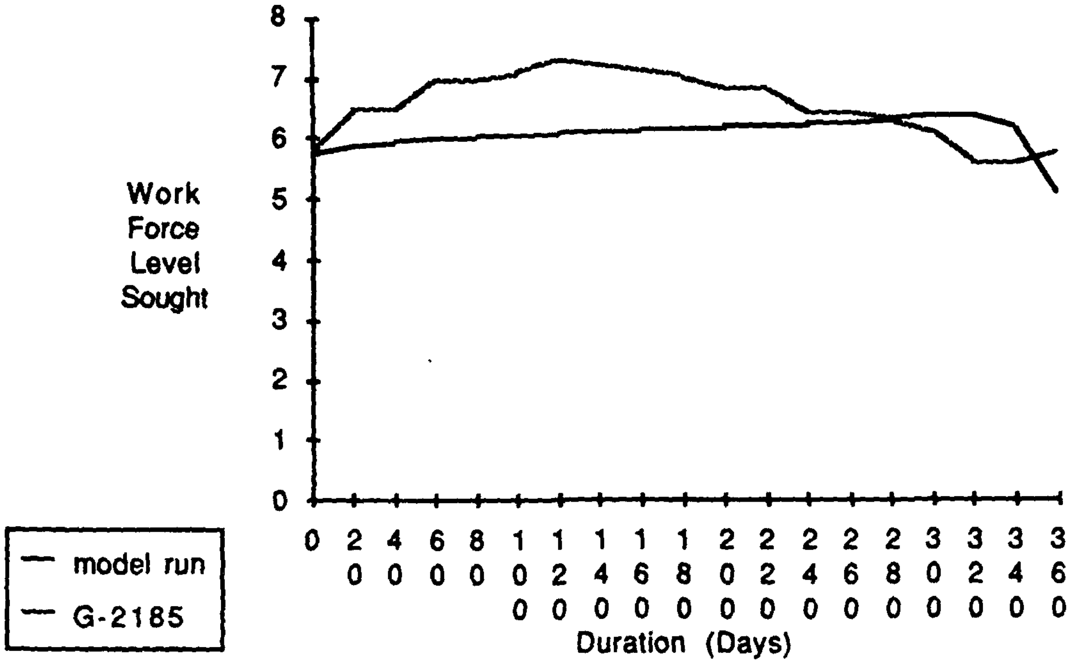

Figure 3-8 "G-2185" WFS Decisions vs. SDM

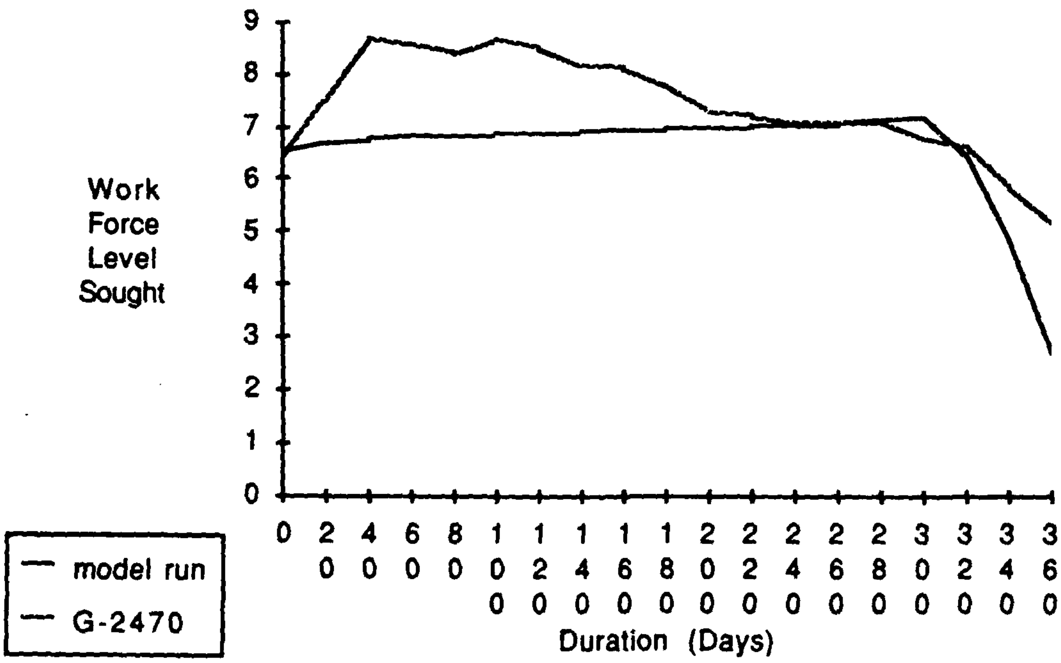

Figure 3-9 "G-2470" WFS Decisions vs. SDM

groups or SDM runs except for the already-explained initial jumps in the groups WFS decisions.

E. CONCLUSIONS

Although not a startling discovery and not the major impetus of this experiment, projects that are under-estimated have been shown to take a significantly longer time to complete. Under-estimation may result in a lower man-day cost if there is no significant schedule pressure towards the end of the development, but the longer duration associated with the project's development may not be worth the man-day cost savings.

The primary objective of this experiment was to compare the groups WFS decisions to those of the SDM running under the exact same conditions. The analysis showed that the experimental groups WFS decisions were significantly different from each other although there were no time nor time and group effects. Compared to the SDM simulation runs, the groups showed the same desired staffing trends and final project durations. The groups did behave in the same manner as the SDM when faced with under-estimation, over-estimation or perfect estimation.

This finding supports the work done by Abdel-Hamid on the utility of using past historical project statistics for cost and schedule estimation [Ref. 3:pp. 1-22]. Showing that real software project managers behave in the same manner as the model under conditions of under-, over- or

perfect estimation proves the usefulness of the SDM for normalizing historical project data and gauging the effectiveness of estimation tools.

IV. "SOCIAL LOAFING" IN SOFTWARE PROJECT MANAGEMENT

A. IMPORTANCE OF THE "SOCIAL LOAFING" PHENOMENON IN SOFTWARE PROJECT MANAGEMENT

The German psychologist Ringelmann conducted an experiment in the 1930's that asked workers to pull as hard as they could on a rope, alone, then with two, three and up to as many as eight other people. In theory, two people should pull twice as hard as one person and eight people should pull eight times harder than a single person. Ringelmann measured the strength of the pulls and discovered an interesting result. The average pull strength with only one worker pulling on the rope was 63 kilograms of pressure. Two workers averaged 59 kilograms per worker. Three workers had an even lower worker average of 54 kilograms. When eight workers were pulling on the rope, the average pull strength per worker was only 32 kilograms of pressure. It seems that in larger groups it is easier to disguise slacking and adopt the mind set to "let the other guy do it." The slacking due to working in a group has been identified as "social loafing." [Ref. 11:p. 126]

Software project management is an endeavor that is in large part performed in groups. The "social loafing" phenomenon, therefore, takes on added importance. Without special attention from senior management, the formation of project management committees or frequent changes in project

leadership can diffuse individual responsibility and lead to "social loafing." In Ringelmann's study, the loss of a few kilograms of pressure due to "social loafing" is interesting but not necessarily critical to the success of the workers. In software project management, the consequences of "social loafing" are profound. The costs for developing software are skyrocketing. Reduced productivity due to the presence of "social loafing" can add a significant cost to an already expensive operation. Senior management must identify and eliminate all controllable factors that reduce the organization's productivity. To counteract the effects of "social loafing," senior management must funnel the social forces present in the organization so that the formation of project committees and changes in project leadership can serve as means of intensifying individual responsibility. [Ref. 11:p. 128]

B. EXPERIMENTAL OBJECTIVE

The objective of this experiment is to determine if software project managers make different project management decisions based on whether they had project responsibility throughout the development phase or whether they assumed project control from another project manager at some time into the development phase. Specifically, the experiment will compare the desired staffing level decisions made by software project managers that have control of a project from start to completion with the staffing decisions of

those that do not assume control until five months (100 work days) into the development phase of an initially scheduled 15 month project.

C. EXPERIMENTAL DESIGN

1. Basic Framework

The experimental objective requires the creation of a project management scenario that can compare the staffing decisions of two groups that assume project management responsibilities at different points in the development phase. A major problem with this simple scenario resides in the fact that each member in the group which assumes responsibility at the start of the development phase will have different project variable values (i.e., experienced work force level, cumulative man-days expended, estimated duration date, percent reported complete, etc.) when the second group is ready to commence its project management responsibilities five months into the development phase. To adequately compare the two groups staffing decisions, though, the experiment must establish a reference point in time from which the two groups can manage the same software project. The current project variable values at this reference point must be identical for the two groups. In other words, the effect of the treatment in the experiment, in this case the different starting points for assuming project management responsibility, must be transparent to the model so that

each subject's behavior is based upon the same starting conditions.

As in the previous experiment, the subjects were designated the "Captains" of the flight simulator. They were to fill the role of software project manager by making the desired staffing level decisions throughout or for the re-mainder of the project's development phase. Regardless of when they started making the desired staffing level decisions, the objective of both groups was to determine a de-sired staffing level for the remainder of the project that they felt provided the best compromise between finishing on an acceptable schedule while avoiding an extensive cost overrun.

The basic framework was to program the experimental model so that the group that assumed responsibility at the start of the development phase would reach the exact same point at which the other group would assume project manager responsibility no matter what staffing decisions the first group made. To do this the experiment had to create a temporary illusion whereby the subjects thought they were managing a project when, in effect, they had absolutely no control over any of the project variables until the second group was ready to begin their project management responsi-bilities. The creation of the illusion involved a number of steps. First, the only project management decision solicited from the subjects by the model was for the desired

staffing level, also known as work force level sought (WFS). WFS is the staffing level that the project manager desires. In the model, as in reality, a project manager does not always get what he/she desires immediately. Factors such as the hiring delay, turnover rate, transfer rate, work force ceiling, and available work force might inhibit attaining the WFS level. Using WFS was important because there were all those uncontrollable factors that could be used to explain the difference between the WFS of the subjects and the model's reported full-time staff.

The model was designed such that for the first 100 days (i.e., until the second group started making project management decisions), the first group's WFS values were ignored by the model. If the subject entered a WFS above the model's generated full-time staff, the model reported the full-time staff and the difference could be attributed to the uncontrollable factors. If the subject reported a WFS below the model's full-time staff, the WFS input would be displayed as the model's full-time staff to prevent the subject from realizing that the model was making staffing decisions that were above the desired staffing level.

Another step in creating the illusion was to limit the number of ignored WFS inputs to five, one for each of the first five months. The WFS input first used by the model was at 100 days into the development phase when the second group started making project decisions. The illusion

was helped by using a project scenario whose reported statistics would not justify any substantial changes in the WFS during those first five months. From the initial project estimates through the reported statistics during the first 80 days of project duration there were no exceptional reports that showed the project falling into any serious schedule delays or cost overruns.

2. Experimental Groups

The two 18 subject groups in this experiment were randomly selected from the three groups in the "Anchoring" experiment. Each 12 subject "Anchoring" group was randomly divided into two, six subject groups through use of a random number table. A single six man group from each "Anchoring" group was combined to form an 18 subject "Social Loafing" group. One group, designated "start," assumed software project management responsibilities at the start of the development phase. The other group, designated "middle," started managing the software project after 100 days of the development phase had elapsed. Appendix K lists the subjects, their group and their final cost and project duration.

3. Documentation

The documentation, listed in Appendices L and M, for each group was slightly different so as to reflect the time period at which the group was to start making desired staffing level decisions. The initial project estimates, staffing variables (i.e., turnover rate, hiring delay,

etc.), organizational history and the short lesson on how to use key pieces of reported information were identical for the two groups. The differences in the documentation were limited to referencing the point in time that the subject was to take control of the project and in emphasizing to the "middle" group that they were taking over a project from a previous project manager. The documentation clearly stated to both groups that they were to determine a desired staffing level for the remainder of the project that they felt provided the best compromise between finishing on an acceptable schedule while avoiding an extensive cost overrun. The importance of meeting the project's initial estimates was stressed to each group.

4. Dynex Gaming Interface Control File

The SDM project used in the experiment was identical for each group. The model had control over all variables until time 100. At this point the model passed control for WFS onto the subject.

The Dynex gaming interface control file was different for the two groups. The control file for the "start" group accepted desired staffing level decisions for the entire project life, but it did nothing with the decisions made from time zero to time 80. The control file for the "middle" group bypassed accepting staffing decisions until it reached time 100. At time 100 it showed the current project statistics, as reported in Appendix N, and solicited

the subject for his desired staffing decision. The current reported statistics at time 100 were identical for each group. The gaming interface control file allowed the subjects in the "start" group to think they were actually making the staffing decisions during the early stages of the development phase.

5. Experiment Execution

The two groups in this experiment were separated during all three experiments. Their seclusion was necessary to prevent them from realizing that they were working on the same projects. Upon completion of experiment two the subjects were given a brief break before commencing the "Social Loafing" experiment. The "start" group was given their documentation to read before they executed the batch file that would begin the experiment. After reading the documentation package, the subjects started the experiment by establishing an initial desired staffing level. While the "middle" group read their documentation during their break, the lab attendants booted the gaming interface control file so that it would reach the point where the current statistics for time 100 appeared. After reading their documentation, the "middle" group made their change to the last project manger's desired staffing level and finished the remainder of the project.

Each subject was required to annotate one of the documentation sheets shown in Appendix O after every desired

staffing level decision. The documentation sheet allowed the subjects to check their progress over time and aided in the analysis of the results.

D. "SOCIAL LOAFING" EXPERIMENT RESULTS AND ANALYSIS

The results for the "Social Loafing" experiment consist of the desired staffing level decisions and the final cost and duration values for 18 subjects in the "start" group and 16 subjects in the "middle" group. The small sample sizes and the large range of final cost and duration values cast a doubt on the normality of the group's results. A formal normality test yielded a p-value of 0.01 that confirmed this doubt and rejected the assumption of normality [Ref. 9:pp. 117-118]. Appendix P contains a listing of the SAS control file used to analyze the experimental data.

Figures 4-1 and 4-2 show a marked difference in the final project totals between the two groups. Assuming non-normality of the data, the nonparametric Wilcoxon Rank Sum test was used to compare the final cost and duration values for the two independent groups. Table 4-1 shows the results of these tests. The null hypothesis for the first test, that the mean final cost for the two groups is equal, was soundly rejected, with a p-value of 0.0006, in favor of the alternate hypothesis that the "start" group expended a significantly higher man-day effort than the "middle" group.

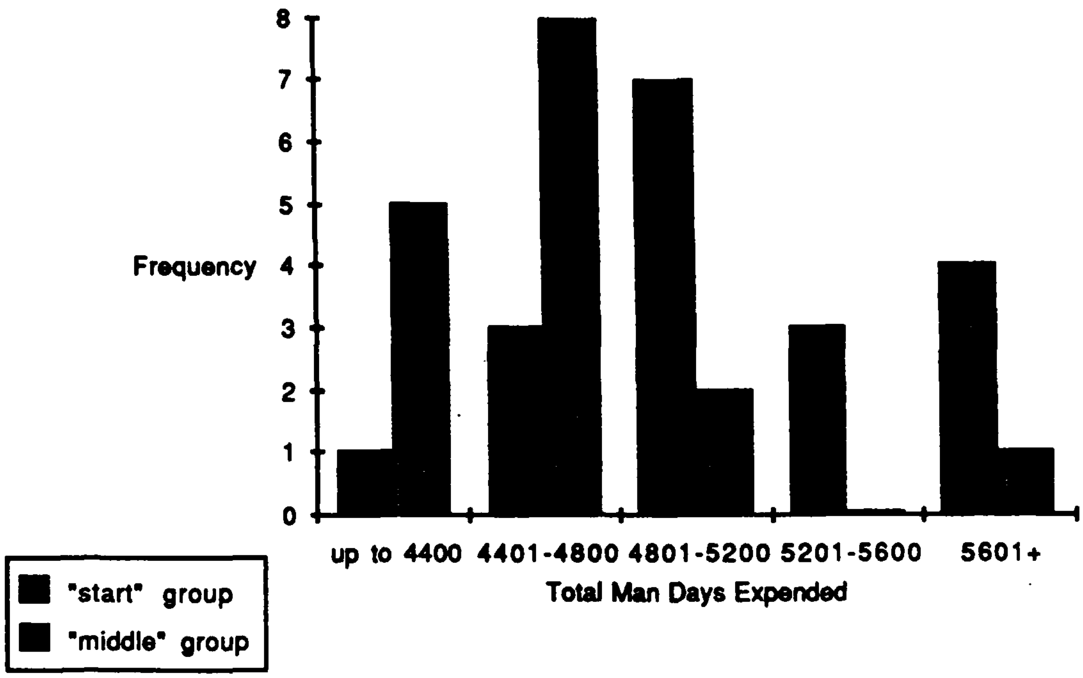

Figure 4-1 Final Cost Comparison

TABLE 4-1

NONPARAMETRIC ANALYSIS OF COST AND DURATION

Group	Mean Project Cost	N	Wilcoxon Scores Sum	Wilcoxon Scores Mean
"Start"	5162	18	414	23.00
"Middle"	4618	16	181	11.31

Wilcoxon 2-Sample Test — S = 181 Z = -3.3988
Prob > Z = 0.0007

Kruskal-Wallis Test — CHISQ = 11.67 DF = 1
Prob > CHISQ = 0.0006

Group	Mean Project Cost	N	Wilcoxon Scores Sum	Wilcoxon Scores Mean
"Start"	414	18	203	11.28
"Middle"	462.5	16	392	24.50

Wilcoxon 2-Sample Test — S = 392 Z = 3.8557
Prob > Z = 0.0001

Kruskal-Wallis — CHISQ = 15.00 DF = 1
Prob > CHISQ = 0.0001

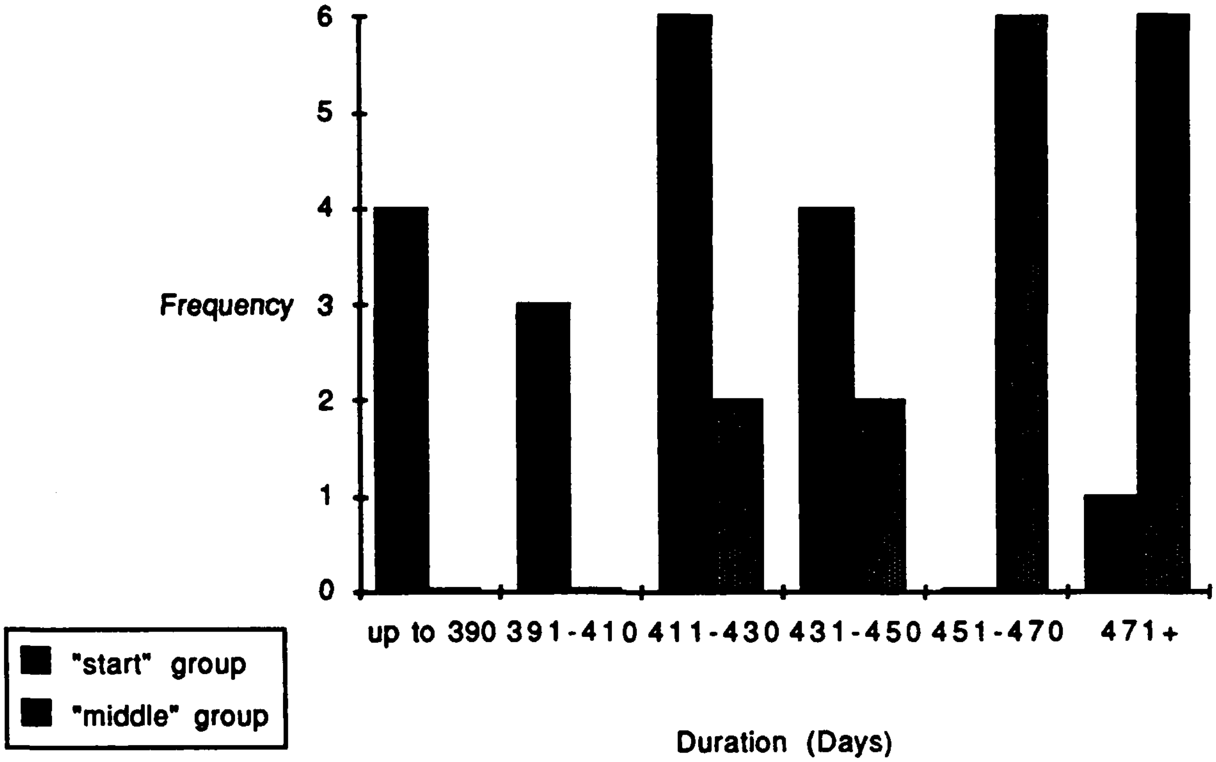

Figure 4-2 Final Project Duration Comparison

The test comparing the final project duration of the two groups resulted in a p-value of 0.0001. The low p-value soundly rejects the null hypothesis in favor of the alternate hypothesis. In this case, the group that assumed project management responsibility at time 100 took a significantly longer time to complete the project.

In addition to analyzing the final project statistics, a comparison of the group's staffing decisions from time 100 through time 400 was made. Figure 4-3 is a plot of the mean WFS decisions made by each group. The "start" group's initial WFS decisions that were ignored by the model are not

shown. The plot of the WFS decisions for each group is terminated once the group's mean final project duration is reached. Stragglers that finished late are not plotted due to the relative distortion their small sample size inflicts on the graph.

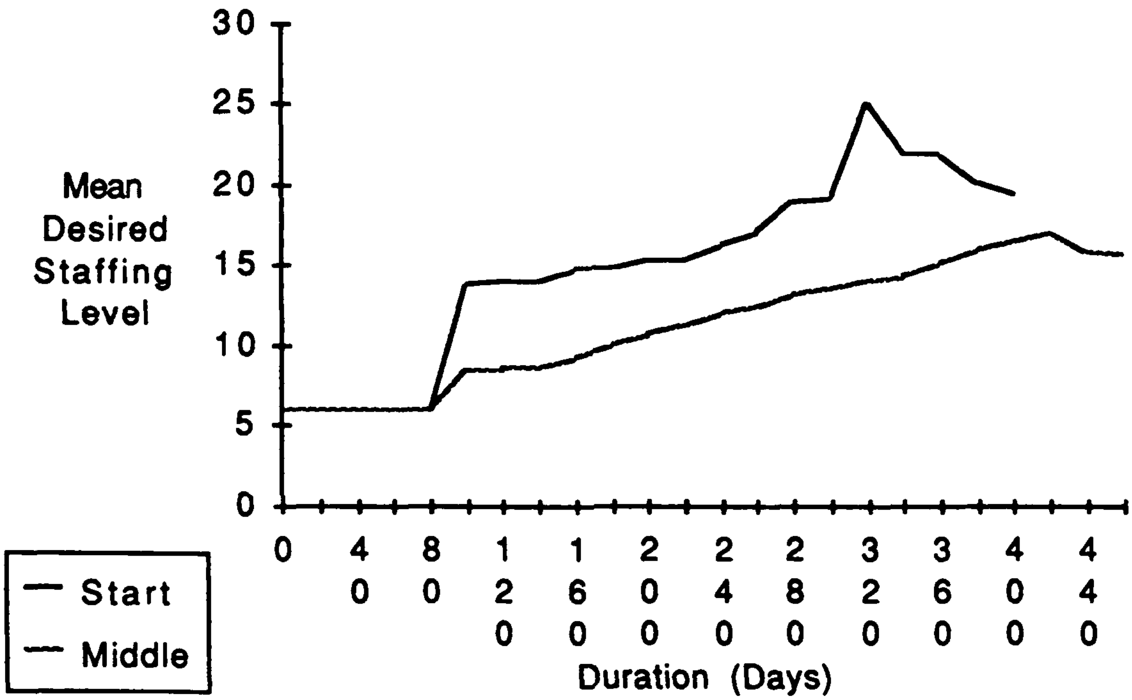

Figure 4-3 Group WFS Decisions

A repeated measures analysis of the data yielded the results in Table 4-2. The first test determines the effect of time on the subject's WFS decisions. The resultant p-value of 0.0013 rejects the null hypothesis of "no time effect." The subject's WFS decisions were influenced by the

point in time at which the WFS decision was made. Referring to Figure 4-3, the rejection of the null hypothesis states that the lines are significantly non-horizontal.

TABLE 4-2

RESULTS OF REPEATED MEASURES TESTS

Test	F-value	Prob > F
No time effect	F(15,15) = 5.27	0.0013
No time and group effect	F(15,15) = 1.31	0.3056
Between subjects effects	F(1,29) = 12.9	0.0012

The next test is for no time and group effect. The result of this test, a p-value of 0.3056, could not reject the null hypothesis. The two group's WFS decisions showed the same trends over time. Again looking back to the graph of the WFS decisions, Figure 4-3, the test states that the lines are not significantly non-parallel.

The last repeated measures test is for the between subjects effects. The p-value of 0.0012 clearly rejects the null hypothesis that the groups made the same WFS decisions over time. In this case, the lines on the graph are not superimposed on each other. The "start" group's mean WFS line is significantly different than the "middle" group's line.

E. CONCLUSIONS

The analysis of the "Social Loafing" experiment yielded significant results. The "start" group showed a deep desire to meet the initial project duration estimate, or to come as close to it as possible, while abandoning a tight control on the project cost. The "middle" group, on the other hand, exhibited an entirely different project management strategy. They kept man-day cost to the minimum while forsaking the project duration. Both groups used the available work force in roughly the same manner (i.e., parallelism and non-horizontalness of the mean WFS lines), but the "start" group used a higher WFS throughout the project life (i.e., the lines were not superimposed) to finish ahead of the "middle" group.

The effect of "social loafing" in this experiment led to an increased project duration and a lower final man-day cost. It appears that project managers that assume responsibility for a project from another manager somewhere during the development phase are profoundly influenced by how the current project statistics at time of relief compare to the initial project estimates. In this experiment (see Appendix N) the project was slightly behind schedule at time 100. Upon seeing that the project was already behind schedule the new project managers started concentrating on cost since they could blame any schedule slippage on the previous project manager. Subject remarks made during the actual

experiment echoed the above observation. The project managers that had responsibility for the project from its inception still concentrated on both cost and schedule throughout the entire development phase.

A post-experiment review of the structure and execution of the experiment identified a possible side effect that may have contributed to the results. The first five WFS decisions for the "start" group were ignored by the model. Although the model was designed so that the subjects should not have wanted to make any drastic increases in the desired staffing level, a subject making that drastic change would not have seen a corresponding jump in the model's full-time work force statistic. Some subjects in the "start" group that did increase their WFS level during the initial five time periods may have felt a lack of control over the work force level through their initial WFS decisions thereby causing them to maintain an artificially high WFS well into the model's responsive time frame. The artificially high WFS level would lead them to a higher cost and lower duration. Table 4-3 shows that the mean WFS decisions for the "start" group were above the reported work force at the next time interval for each of the first five project months.

There is no drastic jump in the mean WFS decision by the subjects in the "start" group. The significance of the difference and its steady rise though, are a point for

TABLE 4-3

"START" GROUP WFS DECISION TIME 0 TIME 80

	Time 0	Time 20	Time 40	Time 60	Time 80
"Start Group" Mean WFS at	9.10	10.16	10.69	11.52	13.08
	Time 20	Time 40	Time 60	Time 80	Time 100
Reported Work Force at	5.50	6.50	7.10	7.50	5.70

concern. The steady rise may indicate that the subjects were either losing faith in the responsiveness of the model or losing faith in the ability of their personnel department to hire additional staff. Only two subjects in this group lied above the mean for each of the first five time intervals. Three subjects showed a steady increase in their work force level throughout the first five time intervals.

There is no way, however, to confirm that the side effect of ignoring the "start" group's first five WFS decisions was present in the experiment. A group debriefing held two days after the experiment did not reveal any overt feelings of the model's non-responsiveness by the "start" subjects. Any future experiments along these lines should consider this side effect in advance and take whatever precautions are necessary to limits its interference.

V. CONCLUSIONS

A. ACCOMPLISHMENTS

The objective of this thesis was to investigate a number of heuristics and biases in the management of software projects. The objective was met through the design, construction and execution of three separate experiments. The experiments used the SDM gaming interface to compare the dynamic decision making behavior of subjects under the effects of different treatments.

The first experiment investigated the "anchoring" phenomenon in software productivity estimation. The second experiment examined the effect of different initial project man-day cost estimates on the subject's desired staffing level decisions. The final experiment focused on the differences in staffing decisions between subjects that "managed" the project throughout the development phase with another group of subjects that assumed project management responsibility five months into the development phase.

B. FUTURE DIRECTION

There are two major paths for further research using the SDM gaming interface to investigate managerial heuristics and biases in software development. The first area involves the replication of the above three experiments using real software project managers as the subjects. Although using

graduate students as surrogates in research studies is useful, tracking the behavior of experienced project managers could provide more significant and noteworthy results.

Anyone replicating these experiments should consider the following lessons learned during the experiment's execution.

- A time slot of at least three hours is needed to run the three experiments successively. Experiments can be run on separate days without much difficulty.

- A few of the subjects focused on the maximum tolerable project duration instead of the initial estimate of project duration as the basis for determining if their project was proceeding on schedule. The current reported project statistics table provided by the SDM gaming interface at each time period should be altered so that the maximum tolerable completion date value is listed under the heading "Initial Estimates" instead of its current position under "Reported Statistics at time ===>." In its present location just below the new estimate of duration it becomes an undesirable reference point for determining schedule slippages. In addition the maximum tolerable completion date does not normally change throughout the project. It should not be listed with the variables that are changing at each time period. This change should be made for all three experiments to eliminate any possible confusion (see Appendix H).

- A post-review of the "anchoring" experiment identified that the SDM gaming interface screen that solicited for the revised estimate of the staff's overall average productivity possibly fostered anchoring. The screen was designed so that the subject could enter a new productivity estimate or just hit "enter" to maintain the old estimate. Changing the wording of the screen to eliminate the phrase, "Press <ENTER> to keep the same productivity estimate," would remove any external "anchoring" influences.

- A work force ceiling, or constraint, should be added to experiment two to prevent subjects from making absurdly high WFS decisions. Two subjects drastically raised their WFS decisions towards the end of development. In real life, a software project manager would encounter

much difficulty trying to raise the work force level 300% towards the end of project development.

- As previously noted in Chapter IV, a further analysis of the effect of ignored WFS decisions for the "start" group in the "social loafing" experiment must be made.

The other area of research involves investigating new managerial heuristics and biases in software project management. The following is a list of possible experimental topics:

- Comparing the final project cost, duration and staffing decisions of subjects that "manage" a project alone with those that manage the project in groups of two or more.

- Comparing the final project cost, duration and staffing decisions of subjects that have a stringent work force ceiling with those that have no imposed work force bounds.

- Determining if tabular reports of current project variable values, as presently used, are superior to plots of the project variables over time. Comparison of final project cost, duration and staffing decisions for three groups that have only plots, only tabular reports or both plots and tabular reports is a viable method for assessing the best output display.

APPENDIX A

IS 4300 STUDENT SURVEY

NAME: ______________________________________
last first m.i.

RANK: ________ **SERVICE:** _________

UNDERGRAD MAJOR: ________________ **COLLEGE GRAD DATE:** ______

NEXT ASSIGNMENT (if known): ________________________________

Brief Job Description: ________________________________

____

PREVIOUS ASSIGNMENTS and EXPERIENCE

Ever employed as a computer programmer? No _____ Yes ____

If employed as programmer, how long (in years) __________
Largest Program worked on (in DSI) __________

Ever employed as a project manager (making personnel decisions and project estimates) for a large project (software or other)?

No _______ Yes ______

If employed as project manager, what was the approximate size of the project in man days or man months? __________
(indicate value and units)

Ever employed as a user or contracting representative responsible for interfacing with or controlling, the money to or the product from a Software Contractor?

No ________ Yes ________

APPENDIX B

"ANCHORING" EXPERIMENT DOCUMENTATION

THE "FLIGHT SIMULATOR" FOR SOFTWARE PROJECT MANAGEMENT

Experiment (1)

INTRODUCTION

The exercise you are about to undertake is similar in many ways to the flight simulators that pilots use to mimic flying an aircraft from takeoff at point A to landing at point B. Instead of flying an aircraft, though, this simulator mimics the life of a real software project from the start of the design phase until the end of testing. In less than an hour you will live through the project's lifecycle. You will be more than an observer. In fact you will play a real role on the project. Your role will not be that of the project manager, but rather of a valuable assistant to the manager (i.e., using the flight simulator analogy again, you can think of your role as that of the flight engineer).

Specifically, your role will be to track the project's progress using a number of reports that will be produced for you at different intervals during the project, and to make your best estimate of the project team's average productivity (in Tasks/man-day). (A task is a unit of work ... you may think of it as a software module 50 lines of code long.) Your estimate will be critically important to the project manager, since he/she will use this information to make the necessary adjustments to the project's staff and schedule. For example, if at some point in the project the amount of work remaining is 100 tasks, and your estimate for the average productivity is 10 tasks/man-day then the project manager will determine that 10 man-days worth of effort is remaining and he/she will use this information to hire or transfer people and/or adjust the schedule.

Your objective is to come up with the best estimate so that your manager can complete the project on budget and on schedule.

THE PROJECT

The project is a real project conducted in a real organization. The organization is on the leading edge in its software engineering practices. It uses a customized version of COCOMO which has been calibrated using the organization's extensive database of historical project data. Further details on the project will be provided later.

YOUR TASK

Your task is to use the reports generated by the project team on resources used to date, work accomplished, and time elapsed to come up with an estimate of the team's overall average productivity that your manager can use in conjunction with other project data to update his/her project plans (e.g., effort remaining, staff needed, scheduled completion date). An example report is attached.

Important things to consider:
- The initial project productivity estimate is derived from an extensive database of historical project statistics that this organization has developed and maintained in the last five years.

- Because software is basically an intangible product during the earlier phases of design and coding, the "% Project Reported Complete" can not be assumed to be totally reliable initially. As one author explained:

> It is essentially impossible for the programmer to estimate the fraction of the program completed. What is 45% of a program? Worse yet, what is 45% of three programs? How is he to guess whether a program is 40% or 50% complete? The easiest way for the programmer to estimate such a figure is to divide the amount of time actually spent on the task to date by the time budgeted for that task. Only when the program is almost finished or when the allocated time budget is almost used up will he be able to recognize that the calculated figure is wrong.

- Factors affecting productivity:

- Workforce mix. When people are hired, they go through an assimilation period (to learn about the specifics of the project) during which they are only half as productive as the "experienced hands" on the project. This assimilation (or training) period is typically one month long.

- Learning. As the project proceeds, expect the productivity of the team as a whole to increase by around 20-30% due to the learning-curve effect.

- Schedule pressure. Productivity can go up or down depending on whether the project falls behind or ahead of schedule (e.g., if people perceive that they are falling behind schedule they may be motivated to work longer hours to bring the project back on track).

REMEMBER: Your objective is to come up with the best estimate for the team's overall average productivity (in tasks/man-day) so that your manager can complete the project on budget and on schedule.

RULES OF THE GAME

- You will be required to provide your estimates for the team's productivity in tasks/man-day four times during the life of the project:
- At the end of the design phase
- At the end of the testing of the first increment
- At the end of the testing of the second increment
- At the end of the project (testing of the final increment)

At each one of these four points, you will be provided with a progress report on the project's status (as reported by the project team members). Give whatever weight you see fit to these reports. You will also have the project's initial estimates (which as mentioned above, are derived from the organization's historical database). Calculate your best estimate for the team's productivity, and input it into the simulator to be used by the project manager in adjusting the project's plans. Also input your estimate on the paper form and submit it to the lab attendant.

- You must work alone.

- YOUR GRADE will be based on how close your estimates are to the project team's true productivity.

SAMPLE PROJECT STATUS REPORT

CURRENT INTERVAL STATISTICS: Elapsed Time = 40

INITIAL ESTIMATES: (These will not change throughout the project)

Project Size	500		Tasks
Man-day Cost	2330.00		Man Days
Project Duration	345		Days
REPORTED STATISTICS at time = = =>	40		Days
% Project Reported Complete	8.43		Percent
Updated Size of Project	500		Tasks
Total Number-Fulltime Equiv Staff	6.5		Fulltime
Effort Expenditures to Date:			
Development Activities	215.98		Man Days
Design and Coding		154.61	Man Days
Rework (i.e. fixing errors)		28.97	Man Days
Quality Assurance		32.40	Man Days
Testing	0.00		Man Days
Total Man Days Expended	215.98		Man Days
New Est of Duration (start-end)	345		Days
Max Tolerable Project Duration	400		Days

Write your new desired staffing level on the documentation sheet provided and press <ENTER>

PRODUCTIVITY ESTIMATE DOCUMENTATION SHEET

NAME: ______________________________

ELAPSED TIME: ______________________ Days
(From the Progress Report on the screen)

YOUR CURRENT :
ESTIMATE OF THE :
OVERALL AVERAGE : Tasks
PRODUCTIVITY : ____________________ Man Day

(YOU MUST TURN THIS SHEET IN TO A LAB ATTENDANT AFTER ENTERING YOUR ESTIMATE AT THE SIMULATION PROMPT!)

PRODUCTIVITY ESTIMATE DOCUMENTATION SHEET
FINAL REPORT

USE THIS FORM AT THE **END** OF THE PROJECT ONLY!

NAME: ______________________________

COMPLETION TIME: ____________________ Days
(From the screen output)

YOUR FINAL :
ASSESSMENT OF THE:
OVERALL AVERAGE : Tasks
PRODUCTIVITY :____________________ Man Day
(to be included in the organization's historical database)

FINAL PROJECT COST: __________________ Man Days
(total man days expended)

FINAL PROJECT SIZE: __________________ Tasks
(perceived size of project)

Explain the critical factors you took into consideration to come up with your estimates during the project:
(Continue on the back if necessary)

____

____

____

____

____

____

____

____

____

____

____

(TURN THIS COMPLETED SHEET IN TO A LAB ATTENDANT AND PRESS Q <ENTER> TO EXIT FROM THE PROGRAM)

MANAGEMENT'S INITIAL PROJECT ESTIMATES

Initial Estimate of Project Size:		**396**	Tasks
Initial Estimate of Overall Average Productivity:		**0.20**	$\frac{\text{Tasks}}{\text{Man Day}}$
Initial Estimate of Project Cost:	$\frac{396}{0.20}$ =	**1,980**	Man days
Initial Estimate of Project Duration:		**320**	Days

There is approximately a two month safety factor applied to the project duration estimate. (i.e., while any schedule slippage is undesirable, a slippage of more than 55 days is untolerable.)

Maximum Tolerable Project Duration:	**375**	Days

In this organization people are typically assigned to more than one project. They may spend anywhere from 20 to 80% of their time on a particular project. FOR CLARITY, THE AVERAGE STAFF SIZE AND SIMULATION OUTPUT WILL BE GIVEN IN FULL-TIME EQUIVALENT EMPLOYEES. One full time equivalent employee is equal to one person who spends 100% of his time on the project or two people that spend 50% of their time on the project.

Average Staff Size:	$\frac{1980}{320}$ =	**6**	Full-time Equivalent Employees

The Project will start with a full time equivalent core team of 1.5 senior designers, and staff up quickly.

MANAGEMENT'S INITIAL PROJECT ESTIMATES

Initial Estimate of Project Size:	**396**	Tasks
Initial Estimate of Overall Average Productivity:	**0.30**	Tasks / Man Day
Initial Estimate of Project Cost:	$\frac{396}{0.30}$ = **1,320**	Man days
Initial Estimate of Project Duration:	**320**	Days

There is approximately a two month safety factor applied to the project duration estimate. (i.e., while any schedule slippage is undesirable, a slippage of more than 55 days is untolerable.)

Maximum Tolerable Project Duration: **375** Days

In this organization people are typically assigned to more than one project. They may spend anywhere from 20 to 80% of their time on a particular project. FOR CLARITY, THE AVERAGE STAFF SIZE AND SIMULATION OUTPUT WILL BE GIVEN IN FULL-TIME EQUIVALENT EMPLOYEES. One full time equivalent employee is equal to one person who spends 100% of his time on the project or two people that spend 50% of their time on the project.

Average Staff Size: $\frac{1320}{320}$ = **4** Full-time Equivalent Employees

The Project will start with a full time equivalent core team of 1.5 senior designers, and staff up quickly.

MANAGEMENT'S INITIAL PROJECT ESTIMATES

Initial Estimate of Project Size:		**396**	Tasks
Initial Estimate of Overall Average Productivity:		**0.45**	$\frac{\text{Tasks}}{\text{Man Day}}$
Initial Estimate of Project Cost:	$\frac{396}{0.45}$ =	**880**	Man days
Initial Estimate of Project Duration:		**320**	Days

There is approximately a two month safety factor applied to the project duration estimate. (i.e., while any schedule slippage is undesirable, a slippage of more than 55 days is untolerable.)

Maximum Tolerable Project Duration: **375** Days

In this organization people are typically assigned to more than one project. They may spend anywhere from 20 to 80% of their time on a particular project. FOR CLARITY, THE AVERAGE STAFF SIZE AND SIMULATION OUTPUT WILL BE GIVEN IN FULL-TIME EQUIVALENT EMPLOYEES. One full time equivalent employee is equal to one person who spends 100% of his time on the project or two people that spend 50% of their time on the project.

Average Staff Size: $\frac{880}{320}$ = **2.75** Full-time Equivalent Employees

The Project will start with a full time equivalent core team of 1.5 senior designers, and staff up quickly.

APPENDIX C

"ANCHORING" EXPERIMENT STUDENT LIST

Low Estimate Group (Ø.2Ø tasks/man-day) n = 9 subjects

	NAME	REASON FOR EXCLUDING OBSERVATIONS
1.	Acton	
2.	Ellis	
3.	Johnson	
4.	Pardini	Misunderstood definition of productivity.
5.	Peterson	
6.	Rodriguez	Misunderstood definition of productivity.
7.	Rouska	
8.	Shuman	
9.	Sweitzer	
1Ø.	Taylor	
11.	Vannortwick	Misunderstood definition of productivity.
12.	Zeiders	

High Estimate Group (Ø.45 tasks/man-day) n = 1Ø subjects

	NAME	REASON FOR EXCLUDING OBSERVATIONS
1.	Beedenbender	
2.	Bell	
3.	Bischoff	
4.	Clemens	
5.	Deleeuw	Did not participate in the experiment.
6.	Drummond	
7.	Garrabrants	
8.	Mostov	
9.	Myers	
1Ø.	Powell	Misunderstood definition of productivity.
11.	Rassatt	
12.	Sablan	

Perfect Estimate Group (Ø.3Ø tasks/man-day) n = 9 subjects

	NAME	REASON FOR EXCLUDING OBSERVATIONS
1.	Ash	
2.	Banham	
3.	Chase	
4.	Kiefer	Misunderstood definition of productivity.
5.	Kirouac	Misunderstood definition of productivity.
6.	Lekey	
7.	Newton	
8.	Santora	Did not participate in the experiment.
9.	Sawyer	
1Ø.	Schwind	
11.	Spaulding	
12.	Triebwasser	

APPENDIX D

"ANCHORING" EXPERIMENT SAS CONTROL FILE

```
CMS FILEDEF ANCHDATA DISK REVANCH TEXT A1;

DATA ANCHOR;
INFILE ANCHDATA;
INPUT
   NAME $ 1-8 ANCGROUP $ 10-14 T0 16-19 T1 21-26 T2 28-33
        T3 35-40 T4 42-47;
   LABEL T1='TIME 100' T2='TIME 200' T3='TIME 300'
         T4='COMPLETION';

PROC SORT DATA=ANCHOR;
  BY ANCGROUP;

*PRELIMINARY STATISTICS *;

PROC MEANS DATA=ANCHOR;
   VAR T1 T2 T3 T4;
   BY ANCGROUP;
   TITLE 'EVALUATION OF EACH GROUP BY TIME INTERVAL';
 RUN;

PROC UNIVARIATE DATA=ANCHOR FREQ;
   VAR T1 T2 T3 T4;
   BY ANCGROUP;
   ID NAME;
   TITLE 'ANCHORING DATA BY TIME INTERVAL';
RUN;

PROC PLOT;
  PLOT T1*ANCGROUP;
  PLOT T2*ANCGROUP;
  PLOT T3*ANCGROUP;
  PLOT T4*ANCGROUP;
RUN;

* REPEATED MEASURES ANALYSIS*;

PROC GLM DATA=ANCHOR;
   CLASS ANCGROUP;
   MODEL T1-T3=ANCGROUP;
   REPEATED TIME / PRINTE SHORT SUMMARY;
   TITLE 'ANCHORING: REPEATED MEASURES TIME 100 TO TIME
         300';
RUN;
```

APPENDIX E

EXPERIMENT "TWO" ROSTERS

GROUP "G-2185" initial man-day estimate of 2185 man-days.

1. Beedenbender
2. Clemen
3. Deleeuw Did not participate
4. Drummond
5. Kirouac
6. Myers
7. Powell
8. Rouska
9. Sablan

GROUP "G-1900" initial man-day estimate of 1900 man-days.

1. Acton
2. Bischoff
3. Ellis
4. Garrabrants
5. Kiefer
6. Mostov
7. Pardini
8. Sweitzer
9. Zeiders

GROUP "G-1460" initial man-day estimate of 1460 man-days.

1. Banham
2. Bell
3. Lekey
4. Rodriguez
5. Schwind
6. Shuman
7. Spaulding
8. Taylor
9. Triebwasser

GROUP "G-2470" initial man-day estimate of 2470 man-days.

1. Ash
2. Chase
3. Johnson
4. Newton
5. Peterson
6. Rassatt
7. Santora Did not participate.
8. Sawyer
9. Vannortwick

APPENDIX F

EXPERIMENT "TWO" DOCUMENTATION

THE "FLIGHT SIMULATOR" FOR SOFTWARE PROJECT MANAGEMENT

Experiment (2)

INTRODUCTION

This exercise utilizes a slightly different version of the software project management "flight simulator" than what you saw in the first exercise. You are no longer just the flight engineer, you have now been promoted to Captain! In this exercise you will again track the project's progress using the available reports but, this time you will be tasked with making the project's staffing decisions. As the project manager, you can hire additional staff or decrease the staffing level as you deem necessary to complete the project. Your objective (like any software project manager) is to manage your resources wisely and efficiently while always aiming to finish the project on schedule (± any safety factor period available).

THE PROJECT

The project is another real project conducted in a second organization which is also on the leading edge in it's software engineering practices and which uses it's own customized version of COCOMO (i.e. calibrated using the organization's database of historical project data). In this organization, project data is collected using Delivered Source Instruction (DSI) units. Some of the project's **initial estimates** are as follows:

- Project Size: 24,400 DSI.
  Schedule Duration: 380 Work Days.
  Acceptable Project Duration: 370 Days to 390 Days.
  Maximum Tolerable Project Duration: 390 Days.

This project is very similar to a project that has just been completed by the organization. You can therefore correctly assume that the above estimates are highly reliable.

YOUR TASK:

Your task is to use the reports generated by the project team at different points in the project on resources used to date, work accomplished, current staffing level and elapsed time, etc., to determine a desired staffing level for the remainder of the project that you feel provides the best compromise between finishing on an acceptable schedule while avoiding an excessive cost overrun.

Important things to consider:

- The hiring delay for new employees can take up to 6 weeks. The assimilation period for a newly hired employee is typically one month long. This is the time needed to train a new employee in the mechanics of the project and bring him/her up to speed. A new employee (i.e. one that is being trained) is only half as productive as an experienced employee.

- The personnel turnover rate is 20% per year.

- As the software project manager, you specify the desired staffing level. The actual staffing level may, of course, be different due to things you can not control such as turnover and lengthy hiring delays.

- The project is initialized with a core team of 1.5 full time equivalent personnel.

- At different points in the project you will be given reported information on the status of the project. Two key pieces of information for this staffing task are: (1) The updated estimate of the total man days (this update can change to reflect the addition of new requirements and/or changes in the estimate of the team's overall productivity); and (2) Effort expenditures to date (also in man days). Subtracting the second from the first yields the "Remaining Effort in man days." Let us say that at some point in the project the "Remaining Effort" is 1000 man days, the remaining time is 100 days and you have 7 full time equivalent employees working. You are, thus, in a position where you have to use your judgement to do one of the following:

1. Stick with the current schedule. If so then you will need a staff size of 1000/100 = 10 full time employees.

2. Stick with your staff size of 7. This means the schedule has to be pushed back. In this case the model will make the appropriate adjustment to the schedule for you. That is extend it to 1000/7 = 143 days.

3. Do a bit of both. That is increase the staff size a bit, say to 8, which will also mean that the schedule will be extended (appropriately by the model) to 1000/8 = 125 days.

RULES OF THE GAME:

- You will be required to provide the new desired staffing level for the project at the beginning of every month (i.e. every 20 work days). The simulation will stop to show current reported statistics and accept a desired staffing level after each 20 work day period. Annotate your desired staffing level on the documentation sheet as well as entering it at the simulation prompt.

- YOU MUST WORK ALONE.

- A lab attendant must verify the final project totals once you have completed the exercise.

MANAGEMENT'S INITIAL PROJECT ESTIMATES

Initial Estimate of Project Size:	24,400	DSI
Initial Estimate of Project Cost:	1460	Man Days
Initial Estimate of Project Duration:	380	Days

Acceptable Duration Range: 370 Days to 390 Days

The Maximum Tolerable Project Duration:	390	Days

MANAGEMENT'S INITIAL PROJECT ESTIMATES

Initial Estimate of Project Size:	24,400	DSI
Initial Estimate of Project Cost:	1900	Man Days
Initial Estimate of Project Duration:	380	Days
Acceptable Duration Range:	370 Days to 390 Days	
The Maximum Tolerable Project Duration:	390	Days

MANAGEMENT'S INITIAL PROJECT ESTIMATES

Initial Estimate of Project Size:	24,400	DSI
Initial Estimate of Project Cost:	2185	Man Days
Initial Estimate of Project Duration:	380	Days
Acceptable Duration Range:	370 Days to 390 Days	
The Maximum Tolerable Project Duration:	390	Days

MANAGEMENT'S INITIAL PROJECT ESTIMATES

Initial Estimate of Project Size:	24,400	DSI
Initial Estimate of Project Cost:	2470	Man Days
Initial Estimate of Project Duration:	380	Days

Acceptable Duration Range: 370 Days to 390 Days

The Maximum Tolerable Project Duration:	390	Days

APPENDIX G

INSTRUCTIONS FOR EXPERIMENT "TWO" GROUPS

Important Points to Remember!!!!!!!!!!
****

- You are not allowed to discuss this exercise with anyone other than a lab attendant. Pleas refrain from discussing this with member in the other class until they have completed the exercise.

- The system will show you the size of the initial core team of senior designers (the full time equivalent number). It will then ask you for your initial desired staffing level. Next it will run through the first simulation time period and show you the current reported statistics. Make your change to the full time equivalent staffing level on the documentation sheet provided after reviewing the report. There is no need to turn in the documentation sheet after each interval.

A LAB ATTENDANT MUST VERIFY YOUR FINAL RESULTS!

- GOOD LUCK! Press <ENTER> to continue.

APPENDIX H

REPORTED PROJECT STATISTICS AT TIME 20
FOR THE "G-2185" EXPERIMENT 2 GROUP

(These statistics are dependent upon an initial WFS decision of 5 at time zero.)

CURRENT INTERVAL STATISTICS: Elapsed Time = 20

INITIAL ESTIMATES: (These will not change throughout the project)			
Project Size	24,400		DSI
Man-day Cost	2,185.00		Man Days
Project Duration	380		Days
REPORTED STATISTICS at time = = =>	20		Days
% Project Reported Complete	1.14		Percent
Updated Size of Project	24,400		DSI
Updated Est. of total Man Days	2,185.00		Man Days
Total Number-Fulltime Equiv Staff	3.2		Fulltime
Effort Expenditures to Date:			
Development Activities	48.35		Man Days
Design and Coding		28.55	Man Days
Rework (i.e. fixing errors)		4.25	Man Days
Quality Assurance		15.55	Man Days
Testing	0.00		Man Days
Total Man Days Expended	48.35		Man Days
New Est of Duration (start-end)	444		Days
Max Tolerable Project Duration	390		Days

Write your new desired staffing level on the documentation sheet provided and press <ENTER>

APPENDIX I

EXPERIMENT "TWO" DOCUMENTATION SHEET

Name: ______________________________

Elapsed Time (days)	Total Man Days Expended (man days)	New Estimate of Project Duration (days)	Staffing Level Sought (FTE Staff)
Initial	Ø	38Ø	
2Ø			
4Ø			
6Ø			
8Ø			
1ØØ			
12Ø			
14Ø			
16Ø			
18Ø			
2ØØ			
22Ø			
24Ø			
26Ø			
28Ø			
3ØØ			
32Ø			
34Ø			
36Ø			
38Ø			

APPENDIX J

EXPERIMENT "TWO" SAS CONTROL FILE

```
CMS FILEDEF EX2FINAL DISK EX2FINAL TEXT A1;
CMS FILEDEF X2240 DISK X2-0-240 TEXT A1;
CMS FILEDEF X2260UP DISK X2-260UP TEXT A1;

DATA X2START;  *WFS DECISIONS TIME 0 TO TIME 240*;
INFILE X2240;
INPUT
NAME $ 1-8 T0 9-12 T20 14-17 T40 19-22 T60 24-27 T80 29-32
     T100 34-37 T120 39-42 T140 44-47 T160 49-52 T180 54-57
     T200 59-62 T220 64-67 T240 69-72;

DATA X2END;    *WFS DECISIONS TIME 260 THROUGH END*;
INFILE X2260UP;
INPUT
NAME $ 1-8 T260 9-12 T280 14-17 T300 19-22 T320 24-27
     T340 29-32 T360 34-37 T380 39-42 T400 44-47 T420 49-52
     T440 54-57 T460 59-62 T480 64-67;

DATA EX2; * GROUP ID, FINAL COST, FINAL DURATION*;
INFILE EX2FINAL;
INPUT
     NAME $ 1-8 EX2GROUP $ 10-15 MD 17-20 DAYS 22-24;
     LABEL MD='TOTAL MANDAYS EXPENDED' DAYS='DURATION';


* PRELIMINARY STATISTICS ON THE SUBJECTS FINAL VALUES *;

PROC SORT DATA=EX2;
  BY EX2GROUP;

PROC MEANS DATA=EX2;
  VAR MD DAYS;
  TITLE 'OVERALL STATS FOR SUBJECTS IN EX2 (ACROSS GROUPS)';
RUN;

PROC MEANS DATA=EX2;
  VAR MD DAYS;
  BY EX2GROUP;
  TITLE 'STATS FOR SUBJECTS WITHIN GROUPS IN EXPERIMENT 2';
RUN;

PROC UNIVARIATE DATA=EX2 FREQ;
  VAR MD DAYS;
  BY EX2GROUP;
  ID NAME;
  TITLE 'EVALUATION OF EACH GROUPS FINAL STATS IN EXP 2';
RUN;
```

```
PROC PLOT;
   PLOT MD*EX2GROUP;
   PLOT DAYS*EX2GROUP;
RUN;


*ANALYSIS OF FINAL DURATION VALUES. *;

* NORMALITY TEST*;

PROC UNIVARIATE DATA=EX2 NORMAL;
  BY GHYP;
  VAR DAYS;
   TITLE 'NORMALITY TEST FOR DURATION TEST';
RUN;

*NONPARAMETRIC ANALYSIS OF DURATION*;

PROC NPAR1WAY DATA=EX2 WILCOXON;
  CLASS GHYP;
  VAR DAYS;
  TITLE 'EXP2: NONPARAMETRIC ANALYSIS OF DURATION DUE TO
         INITIAL EST';
 RUN;


* THIS NEXT SECTION MERGERS THE STAFFING DECISIONS TO THE *
* FINAL STATS AND PERFORMS AN ANALYSIS OF THE STAFFING    *
* DECISIONS WITHIN GROUPS.                                 *;

PROC SORT DATA=X2END;
    BY NAME;

PROC SORT DATA=X2START;
    BY NAME;

PROC SORT DATA=EX2;
    BY NAME;

DATA X2ALL;
    MERGE X2START X2END EX2;
    BY NAME;

PROC SORT DATA=X2ALL;
    BY EX2GROUP;

* PRELIMINARY STATISTICS FOR WFS DECISIONS*;

PROC MEANS DATA=X2ALL;
    VAR T0 T20 T40 T60 T80 T100 T120 T140 T160 T180 T200
        T220 T240 T260 T280 T300 T320 T340 T360 T380 T400
        T420 T440 T460 T480;
```

```
    BY EX2GROUP;
    TITLE 'EVALUATION OF STAFFING DECISIONS BY GROUP';
RUN;

PROC UNIVARIATE DATA=X2ALL FREQ;
    VAR T0 T20 T40 T60 T80 T100 T120 T140 T160 T180 T200
        T220 T240 T260 T280 T300 T320 T340 T360 T380 T400
        T420 T440 T460 T480;
    BY EX2GROUP;
   ID NAME;
   TITLE 'EVALUATION OF STAFFING DECISIONS BY GROUP';
RUN;

PROC PLOT;
     PLOT T0*EX2GROUP;
     PLOT T20*EX2GROUP;
     PLOT T40*EX2GROUP;
     PLOT T60*EX2GROUP;
     PLOT T80*EX2GROUP;
     PLOT T100*EX2GROUP;
     PLOT T120*EX2GROUP;
     PLOT T140*EX2GROUP;
     PLOT T160*EX2GROUP;
     PLOT T180*EX2GROUP;
     PLOT T200*EX2GROUP;
     PLOT T220*EX2GROUP;
     PLOT T240*EX2GROUP;
     PLOT T260*EX2GROUP;
     PLOT T280*EX2GROUP;
     PLOT T300*EX2GROUP;
     PLOT T320*EX2GROUP;
     PLOT T340*EX2GROUP;
     PLOT T360*EX2GROUP;
     PLOT T380*EX2GROUP;
     PLOT T400*EX2GROUP;
     PLOT T420*EX2GROUP;
     PLOT T440*EX2GROUP;
     PLOT T460*EX2GROUP;
     PLOT T480*EX2GROUP;
 RUN;

* REPEATED MEASURES ANALYSIS *;

PROC GLM DATA=X2ALL;
   CLASS EX2GROUP;
   MODEL T0 T20 T40 T60 T80 T100 T120 T140 T160 T180 T200
         T220 T240 T260 T280 T300 T320 T340=EX2GROUP;
   REPEATED TIME / SHORT SUMMARY PRINTE;
   TITLE 'EXP 2:  REPEATED MEASURES TIME 0 TO TIME 340';
RUN;
```

APPENDIX K

"SOCIAL LOAFING" EXPERIMENT STUDENT LIST

Name	Group	Final Man Day Cost	Completion in Days
Acton	Start	4359	500
Ash	Start	4948	420
Bischoff	Start	5981	360
Drummond	Start	5441	416
Johnson	Start	5077	405
Kiefer	Start	5949	395
Kirouac	Start	4639	445
Mostov	Start	4853	425
Newton	Start	4906	420
Peterson	Start	4916	420
Powell	Start	5018	410
Rassatt	Start	4812	435
Rodriguez	Start	5549	390
Rouska	Start	4776	440
Sablan	Start	4712	435
Schwind	Start	5278	415
Shuman	Start	5797	390
Spaulding	Start	5908	330
********************	********	************	**************
Banham	Middle	4571	460
Beedenbender	Middle	4855	435
Bell	Middle	4777	435
Chase	Middle	4315	475
Clemens	Middle	4437	480
Deleeuw	Middle	Did not participate	
Ellis	Middle	4385	465
Garrabrants	Middle	4261	500
Lekey	Middle	4437	470
Myers	Middle	4474	460
Pardini	Middle	4307	465
Santora	Middle	Did not participate	
Sawyer	Middle	4500	475
Sweitzer	Middle	4932	430
Taylor	Middle	4505	470
Triebwasser	Middle	4377	485
Vannortwick	Middle	4484	475
Zeiders	Middle	6274	420

APPENDIX L

"SOCIAL LOAFING" GROUP "START" DOCUMENTATION

THE "FLIGHT SIMULATOR" FOR SOFTWARE PROJECT MANAGEMENT

Experiment (3)

INTRODUCTION

This exercise utilizes the same version of the software project management "flight simulator" that you saw in the previous two exercises. In this exercise (like exercise 2) you will track a project's progress using the available reports and make the project's staffing decisions. This project, however, is larger. As project manager, you can increase or decrease the desired staffing level as you deem necessary to complete the project. Your objective is the same as it was in the past exercise: to manage your resources wisely and efficiently while aiming to finish the project on schedule (+- any safety factor period available).

THE PROJECT

The project is a real project conducted in another organization which uses the most current software engineering practices and it's own customized version of COCOMO (i.e. calibrated using the organization's extensive database of historical project data). Like the organization in exercise two, this organization's data is collected using DSI units. Some of the project's **initial estimates** are as follows:

- Project Size: 42,880 DSI. Like in many other organizations, as the project proceeds new requirements may be added increasing the size (on the average by 50%).

- Schedule Duration: 296 Work Days. The organization has dictated that all projects should be completed within the following range: Initial Schedule Duration +- 20%. For this project the range is 237 days to 355 days. The maximum tolerable project duration from a contractual/ legal point of view is 400 days. The organization highly desires that the project be completed before 355 days due to other software projects needing this staff's resources and for the need to properly package the software project for the user. The significance of the maximum tolerable project duration is that if the project is not completed by 400 days, the organization will be faced with a breach of contract and possible lawsuit from the project's contractee.

YOUR TASK:

Your task is to use the reports generated by the project team at different points in the project on resources used to date, work accomplished, current staffing level and elapsed time, etc., to determine a desired staffing level for the remainder of the project that you feel provides the best compromise between finishing on an acceptable schedule while avoiding an extensive cost overrun.

Important things to consider:

- The initial estimate of project cost is derived from an extensive database of historical project statistics that this organization has developed and maintained. This project is similar to projects that the organization has already completed.

- The hiring delay for new employees can take up to 2 months. The assimilation period for a newly hired employee is typically four months long. This is the time needed to train a new employee in the mechanics of the project and bring him/her up to speed. A new employee (i.e. one that is being trained) is only half as productive as an experienced employee.

- The personnel turnover rate is 30% per year.

- As the software project manager, you specify the desired staffing level. The actual staffing level may, of course, be different due to things you can not control such as turnover and lengthy hiring delays.

- The project is initialized with a core team of 4 full time equivalent personnel.

- At different points in the project you will be given reported information on the status of the project. Two key pieces of information for this staffing task are: (1) The updated estimate of the total man days (this update can change to reflect the addition of new requirements and/or changes in the estimate of the team's overall productivity); and (2) Effort expenditures to date (also in man days). Subtracting the second from the first yields the "Remaining Effort in man days." Let us say that at some point in the project the "Remaining Effort" is 1000 man days, the remaining time is 100 days and you have 7 full time equivalent employees working. You are, thus, in a position where you have to use your judgement to do one of the following:

1. Stick with the current schedule. If so then you will need a staff size of 1000/100 = 10 full time employees.

2. Stick with your staff size of 7. This means the schedule has to be pushed back. In this case the model will make the appropriate adjustment to the schedule for you. That is extend it to 1000/7 = 143 days.

3. Do a bit of both. That is increase the staff size a bit, say to 8, which will also mean that the schedule will be extended (appropriately by the model) to 1000/8 = 125 days.

RULES OF THE GAME:

- You will be required to provide the new staffing level for the project at the beginning of every month (i.e. every 20 work days). The simulation will stop to show current reported statistics and accept a desired staffing level after each 20 work day period. Annotate your desired staffing level on the documentation sheet as well as entering it at the simulation prompt.

- YOU MUST WORK ALONE.

- A lab attendant must verify the final project totals once you have completed the exercise.

MANAGEMENT'S INITIAL PROJECT ESTIMATES

Initial Estimate of Project Size:	42,880	DSI
Initial Estimate of Project Cost:	2,359	Man days
Initial Estimate of Project Duration:	296	Days
Acceptable Project Duration:	237 days to 355 days	
The Maximum Tolerable Project Duration:	400	Days

APPENDIX M

"SOCIAL LOAFING" GROUP "MIDDLE" DOCUMENTATION

THE "FLIGHT SIMULATOR" FOR SOFTWARE PROJECT MANAGEMENT

Experiment (3)

INTRODUCTION

This exercise utilizes the same version of the software project management "flight simulator" that you saw in the previous two exercises. In this exercise (like exercise 2) you will track a project's progress using the available reports and make the project's staffing decisions. The only difference in this experiment is that you have been assigned as project manager 100 work days into the development phase. You are going to take over a project that was initially managed by someone else. As the new project manager, you are free to increase or decrease the desired staffing level as you deem necessary to complete the project in accordance with the initial estimates of project duration and project cost. As in the last exercise, your objective is to manage your resources wisely and efficiently while aiming to finish the project on schedule (+- any safety factor available).

THE PROJECT

The project is a real project conducted in another organization which uses the most current software engineering practices and it's own customized version of COCOMO (i.e. calibrated using the organization's extensive database of historical project data). Like the organization in exercise two, this organization's data is collected using DSI units. Some of the project's **initial estimates** are as:

- Project Size: 42,880 DSI. Like in many other organizations, as the project proceeds new requirements may be added increasing the size (on the average by 50%).

- Schedule Duration: 296 Work Days. The organization has dictated that all projects should be completed within the following range: Initial Schedule Duration +- 20%. For this project the range is 237 days to 355 days. The maximum tolerable project duration from a contractual/ legal point of view is 400 days. The organization highly desires that the project be completed before 355 days due to other software projects needing this staff's resources and for the need to properly package the software project for the user. The significance of the maximum tolerable project duration is that if the project is not completed by 400 days, the organization will be faced with a breach of contract and possible lawsuit from the project's contractee.

The current estimates and project statistics will be available on the screen when you run the simulation.

YOUR TASK:

Your task is to use the reports generated by the project team at different points in the project on resources used to date, work accomplished, current staffing level and elapsed time, etc., to determine a desired staffing level for the remainder of the project that you feel provides the best compromise between finishing on an acceptable schedule while avoiding an excessive cost overrun.

Important things to consider:

- The initial estimate of project cost is derived from an extensive database of historical project statistics that this organization has developed and maintained. This project is similar to projects that the organization has already completed.

- The hiring delay for new employees can take up to 2 months. The assimilation period for a newly hired employee is typically four months long. This is the time needed to train a new employee in the mechanics of the project and bring him/her up to speed. A new employee (i.e. one that is being trained) is only half as productive as an experienced employee.

- The personnel turnover rate is 30% per year.

- As the software project manager, you specify the desired staffing level in full time equivalent employees. The actual staffing level may, of course, be different due to things you can not control such as turnover and lengthy hiring delays.

- At different points in the project you will be given reported information on the status of the project. Two key pieces of information for this staffing task are: (1) The updated estimate of the total man days (this update can change to reflect the addition of new requirements and/or changes in the estimate of the team's overall productivity); and (2) Effort expenditures to date (also in man days). Subtracting the second from the first yields the "Remaining Effort in man days." Let us say that at some point in the project the "Remaining Effort" is 1000 man days, the remaining time is 100 days and you have 7 full time equivalent employees working. You are, thus, in a position where you have to use your judgement to do one of the following:

1. Stick with the current schedule. If so then you will need a staff size of 1000/100 = 10 full time employees.

2. Stick with your staff size of 7. This means the schedule has to be pushed back. In this case the model will make the appropriate adjustment to the schedule for you. That is extend it to 1000/7 = 143 days.

3. Do a bit of both. That is increase the staff size a bit, say to 8, which will also mean that the schedule will be extended (appropriately by the model) to 1000/8 = 125 days.

RULES OF THE GAME:

- You will be required to provide the new desired staffing level for the project at the beginning of each month (i.e. every 20 work days). Initially the output shows the current statistics for an elapsed time of 100 days. You are free to change the current desired staffing level at this point. After every 20 work day period the simulation will stop to show current statistics and accept a new desired staffing level. Remember to annotate your desired staffing level on the documentation sheet as well as entering it at the simulation prompt.

- YOU MUST WORK ALONE.

- A lab attendant must verify the final project totals once you have completed the exercise.

MANAGEMENT'S INITIAL PROJECT ESTIMATES

Initial Estimate of Project Size:	42,880	DSI
Initial Estimate of Project Cost:	2,359	Man days
Initial Estimate of Project Duration:	296	Days
Acceptable Project Duration:	237 days to 355 days	
The Maximum Tolerable Project Duration:	400	Days

APPENDIX N

REPORTED PROJECT STATISTICS AT TIME 100 FOR THE "SOCIAL LOAFING" EXPERIMENT

CURRENT INTERVAL STATISTICS: Elapsed Time = 100

INITIAL ESTIMATES: (These will not change throughout the project)

Project Size	42,880		DSI
Man-day Cost	2,359.00		Man Days
Project Duration	297		Days
REPORTED STATISTICS at time = = =>	100		Days
% Project Reported Complete	22.19		Percent
Updated Size of Project	47,086		DSI
Updated Est. of total Man Days	2,515.33		Man Days
Total Number-Fulltime Equiv Staff	5.7		Fulltime
Effort Expenditures to Date:			
Development Activities	606.15		Man Days
Design and Coding		401.06	Man Days
Rework (i.e. fixing errors)		114.18	Man Days
Quality Assurance		90.92	Man Days
Testing	0.00		Man Days
Total Man Days Expended	606.15		Man Days
New Est of Duration (start-end)	339		Days
Max Tolerable Project Duration	400		Days

Write your new desired staffing level on the documentation sheet provided and press <ENTER>

APPENDIX O

EXPERIMENT THREE DOCUMENTATION SHEET

Name: ______________________________

Elapsed Time (days)	Total Man Days Expended (man days)	New Estimate of Project Duration (days)	Staffing Level Sought (FTE Staff)
Initial	0	296	
20			
40			
60			
80			
100			
120			
140			
160			
180			
200			
220			
240			
260			
280			
300			
320			
340			
360			
380			

EXPERIMENT THREE DOCUMENTATION SHEET

Name: ______________________________

Elapsed Time (days)	Total Man Days Expended (man days)	New Estimate of Project Duration (days)	Staffing Level Sought (FTE Staff)
100			
120			
140			
160			
180			
200			
220			
240			
260			
280			
300			
320			
340			
360			
380			
400			
420			
440			
460			
480			

APPENDIX P

"SOCIAL LOAFING" EXPERIMENT SAS CONTROL FILE

```
CMS FILEDEF SLFINAL DISK SLFINAL TEXT A1;
CMS FILEDEF SL240 DISK SL0-240 TEXT A1;
CMS FILEDEF SL260UP DISK SL250+ TEXT A1;

DATA SLINIT; * WFS DECISIONS TIME 0 TO TIME 240*;
INFILE SL240;
INPUT
NAME $ 1-8 T0 9-12 T20 14-17 T40 19-22 T60 24-27 T80 29-32
     T100 34-37 T120 39-42 T140 44-47 T160 49-52 T180 54-57
     T200 59-62 T220 64-67 T240 69-72;

DATA SLEND; *WFS DECISIONS TIME 260 TO END *;
INFILE SL260UP;
INPUT NAME $ 1-8 T260 9-12 T280 14-17 T300 19-22 T320 24-27
      T340 29-32 T360 34-37 T380 39-42 T400 44-47 T420 49-52
      T440 54-57 T460 59-62 T480 64-67 T500 69-72;

DATA SL; * GROUP ASSIGNMENT, FINAL COST , FINAL DURATION*;
INFILE SLFINAL;
INPUT
     NAME $ 1-8 SLGROUP $ 10-15 MD 17-20 DAYS 22-24;
     LABEL MD='TOTAL MANDAYS EXPENDED' DAYS='DURATION';

PROC SORT DATA=SL;
  BY SLGROUP;


*PRELIMINARY STATISTICS ON COST AND DURATION*;

PROC MEANS DATA=SL;
  VAR MD DAYS;
  TITLE 'OVERALL STATS FOR SUBJECTS IN SOCIAL LOAFING
        (ACROSS GROUPS)';
RUN;

PROC MEANS DATA=SL;
  VAR MD DAYS;
  BY SLGROUP;
  TITLE 'STATS FOR SUBJECTS WITHIN GROUPS IN SOCIAL
        LOAFING';
RUN;

PROC UNIVARIATE DATA=SL FREQ;
  VAR MD DAYS;
  BY SLGROUP;
  ID NAME;
```

```
  TITLE 'EVALUATION OF EACH GROUPS FINAL STATS IN SOCIAL
        LOAFING';
RUN;

PROC PLOT;
   PLOT MD*SLGROUP;
   PLOT DAYS*SLGROUP;
RUN;


*NORMALITY TEST & NONPARAMETRIC ANALYSIS OF COST/DURATION*;

PROC UNIVARIATE DATA=SL NORMAL;
  VAR MD DAYS;
  BY SLGROUP;
  ID NAME;
   TITLE 'SOCIAL LOAFING - TEST FOR NORMALICY';
RUN;

PROC NPAR1WAY DATA=SL WILCOXON;
   CLASS SLGROUP;
   VAR MD;
   TITLE 'SOCIAL LOAFING - NONPARAMETRIC ANALYSIS OF COST':
RUN;

PROC NPAR1WAY DATA=SL WILCOXON;
  CLASS SLGROUP;
  VAR DAYS;
  TITLE 'SOCIAL LOAFING-NONPARAMETRIC ANALYSIS OF DURATION':
RUN;


* THIS NEXT SECTION MERGERS THE STAFFING DECISIONS TO THE *
* FINAL STATS AND PERFORMS A REPEATED MEASURES ANALYSIS OF*
* THE STAFFING DECISIONS WITHIN GROUPS.                   *;

PROC SORT DATA=SLEND;
    BY NAME;

PROC SORT DATA=SLINIT;
    BY NAME;

PROC SORT DATA=SL;
    BY NAME;

DATA SLALL;
    MERGE SLINIT SLEND SL;
    BY NAME;

PROC SORT DATA=SLALL;
    BY SLGROUP;
```

```
*PRELIMINARY STATISTICS FOR WFS DECISIONS*;

PROC MEANS DATA=SLALL;
    VAR T0 T20 T40 T60 T80 T100 T120 T140 T160 T180 T200
        T220 T240 T260 T280 T300 T320 T340 T360 T380 T400
        T420 T440 T460 T480 T500;
    BY SLGROUP;
    TITLE 'EVALUATION OF STAFFING DECISIONS BY GROUP';
RUN;

PROC UNIVARIATE DATA=SLALL FREQ;
    VAR T0 T20 T40 T60 T80 T100 T120 T140 T160 T180 T200
        T220 T240 T260 T280 T300 T320 T340 T360 T380 T400
        T420 T440 T460 T480 T500;
    BY SLGROUP;
    ID NAME;
   TITLE 'EVALUATION OF STAFFING DECISIONS BY GROUP';
RUN;

PROC PLOT;
     PLOT T0*SLGROUP;
     PLOT T20*SLGROUP;
     PLOT T40*SLGROUP;
     PLOT T60*SLGROUP;
     PLOT T80*SLGROUP;
     PLOT T100*SLGROUP;
     PLOT T120*SLGROUP;
     PLOT T140*SLGROUP;
     PLOT T160*SLGROUP;
     PLOT T180*SLGROUP;
     PLOT T200*SLGROUP;
     PLOT T220*SLGROUP;
     PLOT T240*SLGROUP;
     PLOT T260*SLGROUP;
     PLOT T280*SLGROUP;
     PLOT T300*SLGROUP;
     PLOT T320*SLGROUP;
     PLOT T340*SLGROUP;
     PLOT T360*SLGROUP;
     PLOT T380*SLGROUP;
     PLOT T400*SLGROUP;
     PLOT T420*SLGROUP;
     PLOT T440*SLGROUP;
     PLOT T460*SLGROUP;
     PLOT T480*SLGROUP;
 RUN;
```

```
*FINAL REPEATED MEASURES ANALYSIS OF WFS DECISIONS*;

PROC GLM DATA=SLALL;
   CLASS SLGROUP;
   MODEL T100 T120 T140 T160 T180 T200 T220 T240 T260 T280
         T300 T320 T340 T360 T380 T400=SLGROUP;
   REPEATED TIME/SHORT SUMMARY PRINTE;
   TITLE 'SOCIAL LOAFING: REPEATED MEASURES TIME 100 TO TIME
          400';
RUN;
```

LIST OF REFERENCES

1. Schlender, Brenton R., "How to Break the Software Logjam," Fortune, 25 September 1989.

2. Abdel-Hamid, Tarek K., and Madnick, Stuart E., "Lessons Learned from Modeling the Dynamics of Software Development," Center for Information Systems Research Working Paper No. 184, Cambridge, Massachusetts, August 1988.

3. Abdel-Hamid, Tarek K., "On the Utility of Historical Project Statistics for Cost & Schedule Estimation: Results from a Simulation Based Case Study," Paper submitted for publication to IEEE Transactions on Software Engineering, October 1988.

4. Forrester, J.W., Industrial Dynamics, The MIT Press, 1961.

5. Tversky, Amos, and Kahneman, Daniel, "Judgment Under Uncertainty: Heuristics and Biases," Science, 27 September 1974.

6. Remus, William, "Graduate Students as Surrogates for Managers in Experiments on Business Decision Making," Journal of Business Research, V. 14, 1986.

7. Sterman, John D., "Testing Behavioral Simulation Models by Direct Experiment," Sloan School of Management Working Paper D-3783, Cambridge, Massachusetts, February 1986.

8. Defense Documentation Center Report AD F12 052, "A Preliminary Inquiry into the Software Estimation Process," by J.A. Farquhar, August 1970.

9. Little, Ramon C., and Schlotzhauer, Sandra D., SAS System for Elementary Statistical Analysis, SAS Institute Inc., 1987.

10. SAS/STAT Guide for Personal Computers, Version 6 Ed., SAS Institute Inc., 1987.

11. Maital, Shlomo, Minds, Markets, & Money, Basic Books, Inc., 1982.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22304-6145	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93943-5002	2
3.	Dr. Tarek K. Abdel-Hamid, Code AD/Ha Department of Administrative Sciences Naval Postgraduate School Monterey, California 93943-5000	10
4.	Prof. Nancy C. Roberts, Code AS/Rc Department of Administrative Sciences Naval Postgraduate School Monterey, California 93943-5000	2
5.	Daniel Ronan 132 Shenandoah Blvd. Toms River, New Jersey 08753	2
6.	Computer Technology Curricular Office, Code 37 Naval Postgraduate School Monterey, California 93943-5000	1
7.	Commandant (G-PTE-1) U.S. Coast Guard 2100 Second Avenue S.W. Washington, D.C. 20593-0001	1